(12)

EUROPEAN PATENT SPECIFICATION

- (45) Date of publication and mention of the grant of the patent:28.08.1996 Bulletin 1996/35
- (21) Application number: 87905023.5
- (22) Date of filing: 30.06.1987

- (51) Int CI.6: **C07K 14/51**, C12N 15/12, A61K 38/18
- (86) International application number: PCT/US87/01537
- (87) International publication number: WO 88/00205 (14.01.1988 Gazette 1988/02)

(54) NOVEL OSTEOINDUCTIVE COMPOSITIONS

OSTEOINDUKTIVE MITTEL

NOUVELLES COMPOSITIONS OSTEOINDUCTIVES

- (84) Designated Contracting States:

 AT BE CH DE FR GB IT LI LU NL SE
- (30) Priority: 01.07.1986 US 880776 17.12.1986 US 943332 20.03.1987 US 28285 26.03.1987 US 31346
- (43) Date of publication of application: 03.05.1989 Bulletin 1989/18
- (60) Divisional application: 95111771.2
- (73) Proprietor: GENETICS INSTITUTE, INC. Cambridge, Massachusetts 02140 (US)
- (72) Inventors:
 - WANG, Elizabeth, A. Carlisle, MA 01741 (US)
 - WOZNEY, John, M. Hudson, MA 01749 (US)
 - ROSEN, Vicki, A. Boston, MA 02116 (US)

- (74) Representative: VOSSIUS & PARTNER
 Postfach 86 07 67
 81634 München (DE)
- (56) References cited:

EP-A- 0 121 976 EP-A- 0 128 041 EP-A- 0 148 155 EP-A- 0 212 474 WO-A-85/04173 US-A- 4 563 350 US-A- 4 619 989

- PROCEEDINGS OF THE NATL. ACADEMY OF SCIENCES USA, vol. 81, January 1984, Washington, DC (US); URIST, pp. 371-375
- SCIENCE, vol. 220, 13 May 1983, Washington, DC (US); URIST, pp. 680-686
- PROCEEDINGS OF THE NATL. ACADEMY OF SCIENCES USA, vol. 80, November 1983, Washington, DC (US); SAMPATH et al., pp. 6591-6595
- PROCEEDINGS OF THE NATL. ACADEMY OF SCIENCES USA, vol. 78, November 1981, Washington, DC (US); SUGGS et al., pp. 6613-6617

313 578 B1

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

Description

The present invention relates to novel proteins, processes for obtaining them and genes encoding them. These proteins are capable of inducing cartilage and bone formation.

Background

5

15

30

35

40

45

Bone is a highly specialized tissue characterized by an extensive matrix structure formed of fibrous bundles of the protein collagen, and proteoglycans, noncollagenous proteins, lipids and acidic proteins. The processes of bone formation and renewal/repair of bone tissue, which occur continuously throughout life, are performed by specialized cells. Normal embryonic long bone development is preceded by formation of a cartilage model. Bone growth is presumably mediated by "osteoblasts" (bone-forming cells), while remodeling of bone is apparently accomplished by the joint activities of bone-resorbing cells, called "osteoclasts" and osteoblasts. A variety of osteogenic, cartilage-inducing and bone inducing factors have been described. See, e.g. European patent applications 148,155 and 169,016 for discussions thereof.

Brief Description of the Invention

The present invention provides novel proteins in purified form and genes encoding them. Specifically, two of the novel proteins are designated BMP-2 Class I (or BMP-2), and BMP-2 Class II (or BMP-4) wherein BMP is bone morphogenic protein. These proteins are characterized by peptide sequences the same as or substantially homologous to amino acid sequences illustrated in Tables II, III and IV below. They are capable of inducing bone formation at a predetermined site. These bone inductive factors are further characterized by biochemical and biological characteristics including activity at a concentration of 10 to 1000ng/gram of bone in an in vivo rat bone formation assay described below. Proteins of this invention may be encoded by the DNA sequences depicted in the Tables or by sequences capable of hybridizing thereto and coding for polypeptides with bone growth factor biological properties or other variously modified sequences demonstrating such properties.

One of the proteins of the invention is designated BMP-2 Class I (or BMP-2). It is characterized by at least a portion of a peptide sequence the same or substantially the same as that of amino acid #1 through amino acid #396 of Table III which represents the cDNA hBMP-2 Class I. This peptide sequence is encoded by the same or substantially the same DNA sequence, as depicted in nucleotide #356 through nucleotide #1543 of Table III. The human peptide sequence identified in Table III is 396 amino acids in length. hBMP-2 or related bone inductive proteins may also be characterized by at least a portion of this peptide sequence. hBMP-2 Class I is further characterized by the ability to induce bone formation.

The homologous bovine bone inductive protein of the invention designated bBMP-2 Class I (or bBMP-2), has a DNA sequence identified in Table II below which represents the genomic sequence. This bovine DNA sequence has a prospective 129 amino acid coding sequence followed by approximately 205 nucleotides (a presumptive 3' non-coding sequence). bBMP-2, Class I is further characterized by the ability to induce bone formation. A further bone inductive protein composition of the invention is designated BMP-2 Class II or BMP-4. The human protein hBMP-2 Class II (or hBMP-4) is characterized by at least a portion of the same or substantially the same peptide sequence between amido acid #1 through amino acid #408 of Table IV, which represents the cDNA of hBMP-2 Class II. This peptide sequence is encoded by at least a portion of the same or substantially the same DNA sequence as depicted in nucleotide #403 through nucleotide #1626 of Table IV. This factor is further characterized by the ability to induce bone formation.

Another aspect of the invention provides pharmaceutical compositions containing a therapeutically effective amount of one or more bone growth factor polypeptides according to the invention in a pharmaceutically acceptable vehicle. These compositions may further include other therapeutically useful agents. They may also include an appropriate matrix for delivering the proteins to the site of the bone defect and for providing a structure for bone growth. These compositions may be employed in methods for treating a number of bone defects and periodontal disease. These methods, according to the invention, entail administering to a patient needing such bone formation an effective amount of at least one of the novel proteins BMP-2 Class I and BMP-2 Class-II as described herein.

Still a further aspect of the invention are DNA sequences coding on expression for a human or bovine polypeptide having the ability to induce bone formation. Such sequences include the sequence of nucleotides in a 5' to 3' direction illustrated in Tables II, III and IV. Alternatively, a DNA sequence which hybridizes under stringent conditions with the DNA sequences of Tables II, III and IV or a DNA sequence which hybridizes under non-stringent conditions with the illustrated DNA sequences and which codes on expression for a protein having at least one bone growth factor biological property are included in the present invention. Finally, allelic or other variations of the sequences of Tables II, III and IV, whether such nucleotide changes result in changes in the peptide sequence or not, are also included in the present

invention.

10

15

25

Still a further aspect of the invention is a vector containing a DNA sequence as described above in operative association with an expression control sequence. Such vector may be employed in a novel process for producing a bone growth factor polypeptide in which a cell line transformed with a DNA sequence encoding expression of a bone growth factor polypeptide in operative association with an expression control sequence therefor, is cultured. This claimed process may employ a number of known cells as host cells for expression of the polypeptide. Presently preferred cell lines are mammalian cell lines and bacterial cells.

Other aspects and advantages of the present invention will be apparent upon consideration of the following detailed description and preferred embodiments thereof.

Detailed Description of the Invention

The proteins of the present invention are characterized by amino acid sequences or portions thereof the same as or substantially homologous to the sequences shown in Tables II, III and IV. These proteins are also characterized by the ability to induce bone formation.

The bone growth factors provided herein also include factors encoded by the sequences similar to those of Tables II, III and IV, but into which modifications are naturally provided (e.g. allelic variations in the nucleotide sequence which may result in amino acid changes in the polypeptide) or deliberately engineered. For example, synthetic polypeptides may wholly or partially duplicate continuous sequences of the amino acid residues of Tables II, III and IV. These sequences, by virtue of sharing primary, secondary, or tertiary structural and conformational characteristics with bone growth factor polypeptides of Tables II, III and IV may possess bone growth factor biological properties in common therewith. Thus, they may be employed as biologically active substitutes for naturally-occurring bone growth factor polypeptides in therapeutic processes.

Other specific mutations of the sequences of the bone growth factors described herein involve modifications of one or both of the glycosylation sites. The absence of glycosylation or only partial glycosylation results from amino acid substitution or deletion at one or both of the asparagine-linked glycosylation recognition sites present in the sequences of the bone growth factors shown in Tables II, III and IV. The asparagine-linked glycosylation recognition sites comprise tripeptide sequences which are specifically recognized by appropriate cellular glycosylation enzymes. These tripeptide sequences are either asparagine-X-threonine or asparagine-X-serine, where X is usually any amino acid. A variety of amino acid substitutions or deletions at one or both of the first or third amino acid positions of a glycosylation recognition site (and/or amino acid deletion at the second postion) results in non-glycosylation at the modified tripeptide sequence.

The present invention also encompasses the novel DNA sequences, free of association with DNA sequences encoding other proteinaceous materials, and coding on expression for bone growth factors. These DNA sequences include those depicted in Tables II, III and IV in a 5' to 3' direction and those sequences which hybridize under stringent hybridization conditions [see, T. Maniatis et al, Molecular Cloning (A Laboratory Manual), Cold Spring Harbor Laboratory (1982), pages 387 to 389] to the DNA sequences of Tables II, III and IV.

DNA sequences which hybridize to the sequences of Tables II, III and IV under relaxed hybridization conditions and which code on expression for bone growth factors having bone growth factor biological properties also encode bone growth factors of the invention. For example, a DNA sequence which shares regions of significant homology, e. g., sites of glycosylation or disulfide linkages, with the sequences of Tables II, III and IV and encodes a bone growth factor having one or more bone growth factor biological properties clearly encodes a member of this novel family of growth factors, even if such a DNA sequence would not stringently hybridize to the sequence of Tables II, III and IV.

Similarly, DNA sequences which code for bone growth factor polypeptides coded for by the sequences of Tables II, III and IV, but which differ in codon sequence due to the degeneracies of the genetic code or allelic variations (naturally-occurring base changes in the species population which may or may not result in an amino acid change) also encode the novel growth factors described herein. Variations in the DNA sequences of Tables II, III and IV which are caused by point mutations or by induced modifications to enhance the activity, half-life or production of the polypeptides encoded thereby are also encompassed in the invention.

Another aspect of the present invention provides a novel method for producing the novel osteoinductive factors. The method of the present invention involves culturing a suitable cell or cell line, which has been transformed with a DNA sequence coding on expression for a novel bone growth factor polypeptide of the invention, under the control of known regulatory sequences. Suitable cells or cell lines may be mammalian cells, such as Chinese hamster ovary (CHo) cells. The selection of suitable mammalian host cells and methods for transformation, culture, amplification, screening and product production and purification are known in the art. See, e.g., Gething and Sambrook, Nature, 293: 620-625 (1981), or alternatively, Kaufman et al, Mol. Cell. Biol., 5(7):1750-1759 (1985) or Howley et al, U.S. Patent 4,419,446. Another suitable mammalian cell line, which is described in the accompanying examples, is the monkey COS-1 cell line. A similarly useful mammalian cell line is the CV-1 cell line.

Bacterial cells are suitable hosts. For example, the various strains of E. <u>coli</u> (e.g., HB101, MC1061) are well-known as host cells in the field of biotechnology. Various strains of <u>B</u>. <u>subtilis</u>, <u>Pseudomonas</u>, other bacilli and the like may also be employed in this method.

Many strains of yeast cells known to those skilled in the art are also available as host cells for expression of the polypeptides of the present invention. Additionally, where desired, insect cells may be utilized as host cells in the method of the present invention. See, e.g. Miller et al, <u>Genetic Engineering</u>, <u>8</u>:277-298 (Plenum Press 1986) and references cited therein.

Another aspect of the present invention provides vectors for use in the method of expression of these novel osteoinductive polypeptides. Preferably the vectors contain the full novel DNA sequences described above which code for
the novel factors of the invention. Additionally the vectors also contain appropriate expression control sequences permitting expression of the bone inductive protein sequences. Alternatively, vectors incorporating modified sequences
as described above are also embodiments of the present invention and useful in the production of the bone inductive
proteins. The vectors may be employed in the method of transforming cell lines and contain selected regulatory sequences in operative association with the DNA coding sequences of the invention which are capable of directing the
replication and expression thereof in selected host cells. Useful regulatory sequences for such vectors are known to
one of skill in the art and may be selected depending upon the selected host cells. Such selection is routine and does
not form part of the present invention.

A protein of the present invention, which induces bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures. An osteogenic preparation employing one or more of the proteins of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. De novo bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery. An osteogenic factor of the invention may be valuable in the treatment of periodontal disease, and in other tooth repair processes. Such agents may provide an environment to attract bone-forming cells, stimulate growth of bone-forming cells or induce differentiation of progenitors of bone-forming cells. Of course, the proteins of the invention may have other therapeutic uses.

A further aspect of the invention is a therapeutic method and composition for repairing fractures and other conditions related to bone defects or periodontal diseases. Such a composition comprises a therapeutically effective amount of at least one of the bone inductive factor proteins of the invention. The bone inductive factors according to the present invention may be present in a therapeutic composition in admixture with a pharmaceutically acceptable vehicle or matrix. Further therapeutic methods and compositions of the invention comprise a therapeutic amount of a bone inductive factor of the invention with a therapeutic amount of at least one of the other bone inductive factors of the invention. Additionally, the proteins according to the present invention or a combination of the proteins of the present invention may be co-administered with one or more different osteoinductive factors with which they may interact. Further, the bone inductive proteins may be combined with other agents beneficial to the treatment of the bone defect in question. Such agents include, but are not limited to various growth factors. The preparation of such physiologically acceptable protein compositions, having due regard to pH, isotonicity, stability and the like, is within the skill of the art.

30

40

In particular, BMP-2 Class I may be used individually in a pharmaceutical composition. BMP-2 Class I may also be used in combination with one or more of the other proteins of the invention. BMP-2 Class I may be combined with BMP-2 Class II. It may also be combined with BMP-3. Further BMP-2 Class I may be combined with BMP-2 Class II and BMP-3.

BMP-2 Class II may be used individually in pharmaceutical composition. In addition, it may be used in combination with other proteins as identified above. Further it may be used in combination with BMP-3.

The therapeutic method includes locally administering the composition as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone damage. Preferably, the bone growth inductive factor composition would include a matrix capable of delivering the bone inductive factor to the site of bone damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of other materials presently in use for other implanted medical applications.

The choice of material is based on, for example, biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. Similarly, the application of the osteoinductive factors will define the appropriate formulation. Potential matrices for the osteoinductive factors may be biodegradable and chemically defined, such as, but not limited to calcium sulfate, tricalciumphosphate, hydroxyapatite, polylactic acid, polyanhydrides; biodegradable and biologically well defined, such as bone or dermal collagen, other pure proteins or extracellular matrix components; nonbiodegradable and chemically defined, such as sintered hydroxyapatite, bioglass, aluminates, or other ceramics; or combinations of any of the above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalciumphosphate. The bioceramics might also be altered in composition, such as in calcium-aluminate-phos-

phate and processing to alter for example, pore size, particle size, particle shape, and biodegradability.

The dosage regimen will be determined by the attending physician considering various factors which modify the action of such a growth factor, e.g. amount of bone weight desired to be formed, the site of bone damage, the condition of the damaged bone, the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and the composition of BMP's. The addition of other known growth factors, such as IGF 1 (insulin like growth factor 1), to the final composition, may also effect the dosage. Generally, the dosage regimen should be in the range of approximately 10 to 10⁶ nanograms of protein per gram of bone weight desired. Progress can be monitored by periodic assessment of bone growth and/ or repair, e.g. x-rays. Such therapeutic compositions are also presently valuable for veterinary applications due to the lack of species specificity in bone inductive factors. Particularly domestic animals and thoroughbred horses in addition to humans are desired patients for such treatment with the bone inductive factors of the present invention.

The following examples illustrate practice of the present invention in recovering and characterizing the bovine proteins and employing them to recover the human proteins, obtaining the human proteins and in expressing the proteins via recombinant techniques.

EXAMPLE I

15

20

35

40

Isolation of Bovine Bone Inductive Factor

Ground bovine bone powder (20-120 mesh, Helitrex) is prepared according to the procedures of M. R. Urist et al., Proc. Natl Acad. Sci USA, 70:3511 (1973) with elimination of some extraction steps as identified below. Ten kgs of the ground powder is demineralized in succesive changes of 0.6N HCl at 4°C over a 48 hour period with vigorous stirring. The resulting suspension is extracted for 16 hours at 4°C with 50 liters of 2M CaCl₂ and 10mM ethylenediaminetetraacetic acid [EDTA], and followed by extraction for 4 hours in 50 liters of 0.5M EDTA. The residue is washed three times with distilled water before its resuspension in 20 liters of 4M guanidine hydrochloride [GuCl], 20mM Tris (pH 7.4), 1 mM N-ethylmaleimide, 1mM iodoacetamide, 1mM phenylmethylsulfonyl fluoride as described in Clin.Orthop. Rel. Res., 171: 213 (1982). After 16 to 20 hours the supernatant is removed and replaced with another 10 liters of GuCl buffer. The residue is extracted for another 24 hours.

The crude GuCl extracts are combined, concentrated approximately 20 times on a Pellicon apparatus with a 10,000 molecular weight cut-off membrane, and then dialyzed in 50mM Tris, 0.1M NaCl, 6M urea (pH7.2), the starting buffer for the first column. After extensive dialysis the protein is loaded on a 4 liter DEAE cellulose column and the unbound fractions are collected.

The unbound fractions are concentrated and dialyzed against 50mM NaAc, 50mM NaCl (pH 4.6) in 6M urea. The unbound fractions are applied to a carboxymethyl cellulose column. Protein not bound to the column is removed by extensive washing with starting buffer, and the bone inductive factor containing material desorbed from the column by 50mM NaAc, 0.25mM NaCl, 6M urea (pH 4.6). The protein from this step elution is concentrated 20- to 40- fold, then diluted 5 times with 80mM KPO₄, 6M urea (pH6.0). The pH of the solution is adjusted to 6.0 with 500mM K_2HPO_4 . The sample is applied to an hydroxylapatite column (LKB) equilibrated in 80mM KPO₄, 6M urea (pH6.0) and all unbound protein is removed by washing the column with the same buffer. Bone inductive factor activity is eluted with 100mM KPO₄ (pH7.4) and 6M urea.

The protein is concentrated approximately 10 times, and solid NaCl added to a final concentration of 0.15M. This material is applied to a heparin - Sepharose column equilibrated in 50mM KPO₄, 150mM NaCl, 6M urea (pH7.4). After extensive washing of the column with starting buffer, a protein with bone inductive factor activity is eluted by 50mM KPO₄, 700mM NaCl, 6M urea (pH7.4). This fraction is concentrated to a minimum volume, and 0.4ml aliquots are applied to Superose 6 and Superose 12 columns connected in series, equilibrated with 4M GuCl, 20mM Tris (pH7.2) and the columns developed at a flow rate of 0.25ml/min. The protein demonstrating bone inductive factor activity has a relative migration corresponding to approximately 30,000 dalton protein.

The above fractions are pooled, dialyzed against 50mM NaAc, 6M urea (pH4.6), and applied to a Pharmacia MonoS HR column. The column is developed with a gradient to 1.0M NaCl, 50mM NaAc, 6M urea (pH4.6). Active fractions are pooled and brought to pH3.0 with 10% trifluoroacetic acid (TFA). The material is applied to a 0.46 x 25cm Vydac C4 column in 0.1% TFA and the column developed with a gradient to 90% acetonitrile, 0.1% TFA (31.5% acetonitrile, 0.1% TFA to 49.5% acetonitrile, 0.1% TFA in 60 minutes at Iml per minute). Active material is eluted at approximately 40-44% acetonitrile. Aliquots of the appropriate fractions are iodinated by one of the following methods: P. J. McConahey et al, Int. Arch. Allergy, 29:185-189 (1966); A. E. Bolton et al, Biochem J., 133:529 (1973); and D. F. Bowen-Pope, J. Biol. Chem., 237:5161 (1982). The iodinated proteins present in these fractions are analyzed by SDS gel electrophoresis and urea Triton X 100 isoelectric focusing. At this stage, the bone inductive factor is estimated to be approximately 10-50% pure.

EXAMPLE II

Characterization of Bovine Bone Inductive Factor

A. Molecular Weight

Approximately 20ug protein from Example I is lyophilized and redissolved in 1X SDS sample buffer. After 15 minutes of heating at 37°C, the sample is applied to a 15% SDS polyacrylamide gel and then electrophoresed with cooling. The molecular weight is determined relative to prestained molecular weight standards (Bethesda Research Labs). Immediately after completion, the gel lane containing bone inductive factor is sliced into 0.3cm pieces. Each piece is mashed and 1.4ml of 0.1% SDS is added. The samples are shaken gently overnight at room temperature to elute the protein. Each gel slice is desalted to prevent interference in the biological assay. The supernatant from each sample is acidified to pH 3.0 with 10% TFA, filtered through a 0.45 micron membrane and loaded on a 0.46cm x 5cm C4 Vydac column developed with a gradient of 0.1% TFA to 0.1% TFA, 90% CH₃CN. The appropriate bone inductive factor containing fractions are pooled and reconstituted with 20mg rat matrix. In this gel system, the majority of bone inductive factor fractions have the mobility of a protein having a molecular weight of approximately 28,000 - 30,000 daltons.

B. Isoelectric Focusing

20

30

40

The isoelectric point of bone inductive factor activity is determined in a denaturing isoelectric focusing system. The Triton X100 urea gel system (Hoeffer Scientific) is modified as follows: 1) 40% of the ampholytes used are Servalyte 3/10; 60% are Servalyte 7-9. 2) The catholyte used is 40mM NaOH. Approximately 20ug of protein from Example I is lyophilized, dissolved in sample buffer and applied to the isoelectrofocusing gel. The gel is run at 20 watts, 10°C for approximately 3 hours. At completion the lane containing bone inductive factor is sliced into 0.5 cm slices. Each piece is mashed in 1.0ml 6M urea, 5mM Tris (pH 7.8) and the samples agitated at room temperature. The samples are acidified, filtered, desalted and assayed as described above. The major portion of activity as determined in the assay described in Example III migrates in a manner consistent with a pl of 8.8 - 9.2.

C. Subunit Characterization

The subunit composition of bone inductive factor is also determined. Pure bone inductive factor is isolated from a preparative 15% SDS gel as described above. A portion of the sample is then reduced with 5mM DTT in sample buffer and re-electrophoresed on a 15% SDS gel. The approximately 30kd protein yields two major bands at approximately 20kd and 18kd, as well as a minor band at 30kd. The broadness of the two bands indicates heterogeneity caused most probably by glycosylation, other post translational modification, proteolytic degradation or carbamylation.

EXAMPLE III

Biological Activity of Bone Inductive Factor

A rat bone formation assay according to the general procedure of Sampath and Reddi, <u>Proc. Natl. Acad. Sci. U. S.A.</u>, 80:6591-6595 (1983) is used to evaluate the osteogenic activity of the bovine bone inductive factor of the present invention obtained in Example I. This assay can also be used to evaluate bone inductive factors of other species. The ethanol precipitation step is replaced by dialyzing the fraction to be assayed against water. The solution or suspension is then redissolved in a volatile solvent, e.g. 0.1 - 0.2 % TFA, and the resulting solution added to 20mg of rat matrix. This material is frozen and lyophilized and the resulting powder enclosed in #5 gelatin capsules. The capsules are implanted subcutaneously in the abdominal thoracic area of 21 - 49 day old male long Evans rats. The implants are removed after 7 - 14 days. Half of each implant is used for alkaline phosphatase analysis [See, A. H. Reddi et al., <u>Proc. Natl. Acad. Sci.</u>, 69:1601 (1972)] and half is fixed and processed for histological analysis. Routinely, 1µm glycolmethacrylate sections are stained with Von Kossa and acid fuchsin to detect new bone mineral. Alkaline phosphatase, an enzyme produced by chondroblasts and osteoblasts in the process of matrix formation, is also measured. New cartilage and bone formation often correlates with alkaline phosphatase levels. Table I below illustrates the dose response of the rat matrix samples including a control not treated with bone inductive factor.

TABLE 1

Protein* Implanted μg	Cartilage	Alk. Phos.u/l
7.5	2	Not done
2.5	3	445.7
0.83	3	77.4
0.28	0	32.5
0.00	0	31.0

*At this stage the bone inductive factor is approximately 10-15% pure.

The bone or cartilage formed is physically confined to the space occupied by the matrix. Samples are also analyzed by SDS gel electrophoresis and isoelectric focusing as described above, followed by autoradiography. Analysis reveals a correlation of activity with protein bands at 28 - 30kd and a pl 9.0. An extinction coefficient of 1 OD/mg-cm is used as an estimate for protein and approximating the purity of bone inductive factor in a particular fraction. In the <u>in vivo</u> rat bone formation assays on dilutions as described above, the protein is active <u>in vivo</u> at 10 to 200ng protein/gram bone to probably greater than 1µg protein/gram bone.

EXAMPLE IV

5

10

20

30

35

50

55

Bovine Bone Inductive Factor Protein Composition

The protein composition of Example IIA of molecular weight 28 - 30kd is reduced as described in Example IIC and digested with trypsin. Eight tryptic fragments are isolated by standard procedures having the following amino acid sequences:

Fragment 1: AAFLGDIALDEEDLG

Fragment 2: A F Q V Q Q A A D L

Fragment 3: N Y Q D M V V E G

Fragment 4: STPAQDVSR

Fragment 5: NQEALR

Fragment 6: LSEPDPSHTLEE

Fragment 7: F D A Y Y

Fragment 8: L K P S N ? A T I Q S I V E

A less highly purified preparation of protein from bovine bone is prepared according to a purification scheme similar to that described in Example I. The purification basically varies from that previously described by omission of the DE-52 column, the CM cellulose column and the mono s column, as well as a reversal in the order of the hydroxylapatite and heparin sepharose columns. Briefly, the concentrated crude 4 M extract is brought to 85% final concentration of ethanol at 4 degrees. The mixture is then centrifuged, and the precipitate redissolved in 50 mM Tris, 0.15 M NaCl, 6.0 M urea. This material is then fractionated on Heparin Sepharose as described. The Heparin bound material is fractionated on hydroxyapatite as described. The active fractions are pooled, concentrated, and fractionated on a high resolution gel filtration (TSK 30000 in 6 M guanidinium chloride, 50 mM Tris, pH 7.2). The active fractions are pooled, dialyzed against 0.1% TFA, and then fractionated on a C4 Vydac reverse phase column as described. The preparation is reduced and electrophoresed on an acrylamide gel. The protein corresponding to the 18K band is eluted and digested with trypsin. Tryptic fragments are isolated having the following amino acid sequences:

Fragment 9: SLKPSNHATIOS? V

Fragment 10: S F D A Y Y C S ? A

Fragment 11: V Y P N M T V E S C A

Fragment 12: V D F A D I ? W

Tryptic Fragments 7 and 8 are noted to be substantially the same as Fragments 10 and 9, respectively.

A. bBMF-2

5

15

Two probes consisting of pools of oligonucleotides are designed on the basis of the amino acid sequence of Fragment 3 and synthesized on an automated DNA synthesizer as described above.

Probe #1: A C N A C C A T [A/G] T C [T/C] T G [A/G] A T Probe #2: C A [A/G] G A [T/C] A T G G T N G T N G A

These probes are radioactively labeled and employed to screen the bovine genomic library constructed as follows:
Bovine liver DNA is partially digested with the restriction endonuclease enzyme Sau 3A and sedimented through a sucrose gradient. Size fractionated DNA in the range of 15-30kb is then ligated to the lambda J1 BamH1 arms vector [Frischauf et al, J. Mol. Biol., 170:827-842 (1983) Mullins et al., Nature 308: 856-858 (1984)]. The library is plated at 8000 recombinants per plate. Duplicate nitrocellulose replicas of the plaques are made and amplified according to a modification of the procedure of Woo et al, Proc. Natl. Acad. Sci. USA, 75:3688-91 (1978).

The radioactively labelled 17-mer Probe #1 is hybridized to the set of filters according to the following method:

The probe is kinased and hybridized to the other set of filters in 3M tetramethylammonium chloride (TMAC), 0.1M sodium phosphate pH6.5, 1mM EDTA, 5X Denhardts, 0.6% SDS, 100ug/ml salmon sperm DNA at 48 degrees C, and washed in 3M TMAC, 50mM Tris pH8.0 at 50 degrees C. These conditions minimize the detection of mismatches to the probe pool [see, Wood et al, <u>Proc. Natl. Acad. Sci. U.S.A.</u>, 82:1585-1588 (1985)]. 400,000 recombinants are screened by this procedure. One duplicate positive is plaque purified and the DNA is isolated from a plate lysate of the recombinant bacteriophage designated lambda bP-21. Bacteriophage bP-21 was deposited with the ATCC under accession number ATCC 40310 on March 6, 1987. The bP-21 clone encodes the bovine growth factor designated bBMP-2.

The oligonucleotide hybridizing region of this bBMP-2 clone is localized to an approximately 1.2 kb Sac I restriction fragment which is subcloned into M13 and sequenced by standard techniques. The partial DNA sequence and derived amino acid sequence of this Sac I fragment and the contiguous Hind III-Sac I restriction fragment of bP-21 are shown below in Table II. The bBMP-2 peptide sequence from this clone is 129 amino acids in length and is encoded by the DNA sequence from nucleotide #1 through nucleotide #387. The amino acid sequence corresponding to the tryptic fragment isolated from the bovine bone 28 to 30kd material is underlined in Table II. The underlined portion of the sequence corresponds to tryptic Fragment 3 above from which the oligonucleotide probes for bBMP-2 are designed. The predicted amino acid sequence indicates that tryptic Fragment 3 is preceded by a basic residue (K) as expected considering the specificity of trypsin. The arginine residue encoded by the CGT triplet is presumed to be the carboxy-terminus of the protein based on the presence of a stop codon (TAG) adjacent to it.

35

30

40

45

50

TABLE II

5	(1) GGC G		GAT D	GGG G	15 AAA K	GGA G	CAC H	CCT P	CTC L	30 CAC H	AGA R	AGA R	GAA E	X AAG	45 CGG R
			AAA K	CAC H	60 AAA K	CAG Q	CGG R	AAA K	CGC R	75 CTC L	aag K	TCC S	AGC S	TGT C	90 AAG K
10	aga R		CCT P	TTA L	105 TAT Y			TTC F		120 GAT D	GTG V	GGG G	TGG W	AAT N	135 GAC D
15		ATC I	GTT V		150 CCG P	CCG P	GGG G	TAT Y	CAT H	165 GCC A	TTT F	TAC Y	TGC C	CAT	180 GGG G
			CCT P	TTT F	195 CCC P	CTG L	GCC A	GAT D	CAC	210 CTT L	AAC	TCC S	ACG T	AAT N	225 CAT H
20	GCC A	ATT I	V V	CAA Q	240 ACT T	CTG	GTC V	AAC N	TCA S	255 GTT V	AAC	TCT S	aag K	ATT I	270 CCC P
 25		GCA A			-385 GTC V	CCA	ACA	GAG E	CTC L	300 AGC S	GCC A	ATC I	TCC S	ATG M	315 CTG L
		CTT L	GAT D	GAG E	0EE AAT N	GAG	aag K	GTG V	GTA V	345 TTA L	AAG	AAC N		CAG O	360 GAC D
30	ATG M					TGT	GGG	TGT C	CGT	9) TAG	: CACA	397 GCA /	AAAT	4 AAAA	07 TA
	TAA		417 ATA 1	TATA:	4: TATA	27 TA T	raga:	43 AAAA	7 C AG	CAAA	447 AAAA	TCA	AGTT	457 GAC	•
35	ACT:	TAA!	467 TAT '	TTCC	4°	77 GA A	GACT	48 TAT	7 F TA:	rgga	497 ATGG	AAT	GGAG	507 AAA	
	AAG		517 ACA	CAGC:		27 PT G	AAAA	53 CTAT	7 A TT:	TATA!	547 ICTA		Aaaa		
40	GTT	-	567 AAA	CAAA:		77 TT A	ATCA	58 GAGA		ATT					

EXAMPLE V

45

Human Bone Inductive Factors

A. hBMP-2: Class I and II

The HindIII-SacI bovine genomic bBMP-2 fragment described in Example IV A. is subcloned into an M13 vector. A ³²p-labeled single-stranded DNA probe is made from a template preparation of this subclone. This probe is used to screen polyadenylated RNAs from various cell and tissue sources.

Polyadenylated RNAs from various cell and tissue sources are electrophoresed on formaldehyde-agarose gels and transferred to nitrocellulose by the method of Toole et al., <u>supra</u>. The probe is then hybridized to the nitrocellulose blot in 50% formamide, 5 X SSC, 0.1% SDS, 40 mM sodium phosphate pH6.5, 100 µg/ml denatured salmon sperm DNA, and 5 mM vanadyl ribonucleosides at 42° C overnight and washed at 65° C in 0.2 X SSC, 0.1% SDS. Following autoradiography, a hybridizing band corresponding to an mRNA species of approximately 3.8 kb is detected in the lane containing RNA from the human cell line U-2 OS. The HindIII-SacI fragment is labeled with ³²P by nick translation and

used to screen the nitrocellulose filter replicas of a U-2 OS cDNA library by hybridization in standard hybridization buffer at 65° overnight followed by washing in 1 X SSC, 0.1% SDS at 65°.

This library was constructed by synthesizing cDNA from U-2 OS polyadenylated RNA and cloning into lambda gt10 by established techniques (Toole et al., supra). Twelve duplicate positive clones are picked and replated for secondaries. Duplicate nitrocellulose replicas are made of the secondary plates and both sets hybridized to the bovine genomic probe as the primary screening was performed. One set of filters is then washed in 1 X SSC, 0.1% SDS; the other in 0.1 X SSC, 0.1% SDS at 65°.

Two classes of hBMP-2 cDNA clones are evident based on strong (4 recombinants) or weak (7 recombinants) hybridization signals under the more stringent washing conditions (0.1 X SSC, 0.1% SDS). All 11 recombinant bacteriophages are plaque purified, small scale DNA preparations made from plate lysates of each, and the inserts subcloned into pSP65 and into M13 for sequence analysis. Sequence analysis of the strongly hybridizing clones designated hBHP-2 Class I (also known as BMP-2) indicates that they have extensive sequence homology with the sequence given in Table II. These clones are therefore cDNA encoding the human equivalent of the protein encoded by the bBMP-2 gene whose partial sequence is given in Table II. Sequence analysis of the weakly hybridizing recombinants designated hBMP-2 Class II (also known as BMP-4) indicates that they are also quite homologous with the sequence given in Table II at the 3' end of their coding regions, but less so in the more 5' regions. Thus they encode a human protein of similar, though not identical, structure to that above.

Full length hBMP-2 Class I cDNA clones are obtained in the following manner. The 1.5 kb insert of one of the Class Il subclones (II-10-1) is isolated and radioactively labeled by nick-translation. One set of the nitrocellulose replicas of the U-2 OS cDNA library screened above (50 filters, corresponding to 1,000,000 recombinant bacteriophage) is rehybridized with this probe under stringent conditions (hybridization at 65° in standard hybridization buffer; washing at 65° in 0.2 X SSC, 0.1% SDS). All recombinants which hybridize to the bovine genomic probe which do not hybridize to the Class II probe are picked and plaque purified (10 recombinants). Plate stocks are made and small scale bacteriophage DNA preparations made. After subcloning into M13, sequence analysis indicates that 4 of these represent clones which overlap the original Class I clone. One of these, lambda U2OS-39, contains an approximately 1.5 kb insert and was deposited with the ATCC on June 16, 1987 under accession number 40345. The partial DNA sequence (compiled from lambda U2OS-39 and several other hBMP-2 Class I cDNA recombinants) and derived amino acid sequence are shown below in Table III. Lambda U2OS-39 is expected to contain all of the nucleotide sequence necessary to encode the entire human counterpart of the protein BMP-2 Class I encoded by the bovine gene segment whose partial sequence is presented in Table II. This human cDNA hBMP-2 Class I contains an open reading frame of 1188 bp, encoding a protein of 396 amino acids. This protein of 396 amino acids has a molecular weight of 45kd based on this amino acid sequence. It is contemplated that this sequence represents the primary translation product. The protein is preceded by a 5' untranslated region of 342 bp with stop codons in all frames. The 13 bp region preceding this 5' untranslated region represents a linker used in the cDNA cloning procedure.

25

30

35

40

45

50

55

TABLE III

5																				
	GIO	GACI	10 CIA	GAGI	GIGI	20 GI (AGCA		0 G CI	GGGG	40 ACTT		GAAC	50	CAGG		60 AT 2	AACIT	70 GCGCA	
10	\sim	ריים	80 בחייוי	\sim		90 ∝ 1	TTTCC	10		~~~	110			120	~~~1		30	~~~	140	
	-	-CACI	110	u.	196316	CC 1	.1160	LLL'A	له لا	CALC	CIGO	110	LCCA	ICI.	ССБА	GCC	LA (XICC	
	۵ζ۳		150	CTTTC:		.60		17		•	180			190	2020		00	20000	210	
15	1101			CIIG		·	(CI GP	·	C 16	,1100	الملاحث	616	ereren	13/113	AGAC	1603		GCCCGC	ACC	
			220			30	٠	24			250			260		2	70		280	
20	GGG	ALAA	GGA	GGAG	GCAA	AG A	AAAG	GAAC	G GA	CATT	OGGT	œı	TGCG	CCA.	GGIC	CITI	GA (CAGA	TITT	
20			290		3	00		23	o.		220			226		_	40			
	TCC			œcr			TGGA	021G 31		2000	GIGC			330 ACG	GACT		40 TC 1	rccia;	350 VAGGT	
25		(1)			3	70				31	85				4	00			
	CGA	∝ A M	rg g er v	TG G	∝ G la G	GG A	∞ o	ec u	GT C YS L	TT C eu L	IA G	OG T	TG C eu L	rg c eu I	TT C eu P	∞ c	AG C	TC /al		
				415					430					445						
30	CTC	CIG	GCC	GCC Glv	GCG Ala	GCT Ala	GGC Gly	CTC	GTT Val	CCG Pro	GAG	CTG	GGC	OGC	AGG	AAG	TTC	GCG Ala		
			•	•			2						CLY	129	my		1.11	·		
	460 GCG		TOG	TCG	ccc	475	α	מיאוי	тv	CAG	490	יוריאוו	CNC	CNC	CIV.	505	700	GAG		
35	Ala	Ala	Ser	Ser	Gly	Arg	Pro	Ser	Ser	Gln	Pro	Ser	Asp	Glu	Val	Leu	Ser	Glu		
			520					535					550					565		
	TTC	GAG	TIG	∞	CIG	CTC	AGC	ATG	TTC	GGC	CIG	AAA	CAG	AGA	∞	ACC	∞	AGC		
	.me	GIU	Leu	Arg	Leu	Leu	Ser	MET	Phe	Gly	Leu	Lys	Gln	Arg	Pro	Thr	Pro	Ser		
40					580					595					610					
	AGG	GAC	GCC	GTG	GTG	∞	∞	TAC	ATG	CIA	GAC	CIG	TAT	œc	AGG	CAC	TOG	GGT		
	Arg		wra	Val	vai	PIO	Pro	ıyr	WET.	Leu	Asp	Leu	Tyr	Arg	Arg	His	Ser	Gly		
	03.0	625					640					655					670			
45	Gln	Pro	Gly	Ser	Pro	GCC Ala	CCA Pro	GAC Asp	CAC His	Arq	TTG	GAG Glu	AGG	GCA Ala	GCC Ala	AGC Ser	Arq	GCC		
			_	685				•	700					715			- 3			
	AAC	ACT	GTG		AGC	TIC	CAC	CAT		GAA	TCT	TIG	GAA		CTA	CCA	GAA	ACG		
50	Asn	Thr	Val	Arg	Ser	Phe	His	His	Glu	Glu	Ser	Leu	Glu	Glu	Leu	Pro	Glu	Thr		

	730 AGI Ser	' GGG	AAA Lys	ACZ Thr	ACC Thr	745 C CGC Arc	AGA	TTC	TTC Phe	TTI	760 CAA CAST	TTA	AGI Ser	TCI Ser	ATC	775 000 Pro	: ACC	GAG Glu
.	GAG Glu	TTT Phe	790 ATC	ACC	TCA Ser	GCA Ala	GAG Glu	805 CIT Leu	CAG	GII Val	TTC Phe	C CGA	820 GAA Glu	CAG	ATG MET	CAA Gln	GAI Asp	835 CCT Ala
10	TTA Leu	GGA Gly	AAC	AA1 Ast	850 AGC Ser	AGI	TIC Phe	CAT His	CAC His	865 CGA Arg	ATI	'AAT Asn	ATT	TAT Tyr	880 GAA Glu	ATC	ATA Ile	AAA : Lys
15	CCT Pro	895 GCA Ala	ACA	GCC Ala	AAC Asn	TOG Ser	910 AAA Lys	TTC	Pro	GIG Val	ACC Thr	925 AGI Ser	CIT	TIG	GAC Asp	ACC Thr	940 AGG Arg	TTG Leu
· 20	Val	Asn	CAG Gln	955 AAT Asn	'GCA Ala	Ser	Arg	TGG	970 GAA Glu	AGI	TTI Phe	GAT Asp	GIC Val	985 ACC Thr	∞	GCT Ala	GIG Val	ATG MET
	100 Arg	IGG Trp	Thr	Ala	. CAG	Glv	CAC His	Ala	Asn	CAT	Glv	TTC	Val	1721	GAA	GIG Val	GCC	CAC His
25	TTG Leu	GAG	1060 GAG Glu	AAA Lys	GLD	Gly	GIC	1075 TCC Ser	'Lys	Arg	His	GIT	1090 AGG Arg	ATA He	Ser	AGG Airg	יויטד	1105 TIG Leu
30	HIS	CAA Gln Ll65	GAT Asp	GAA	1120 CAC His	AGC Ser	'Trp	TCA Ser	CAG	1135 ATA Ile	AGG Arg	Pro	TTG Leu	CTA	II50 GTA Val	ACT Thr	TTT Phe	GGC Gly
35	CAT	GAT	GIY	Lys	GGG	CAT	L180 CCT Pro	Leu	His	aaa Lys	AGA	1195 GAA Glu	Lys	Arg	CAA Gln	GCC	1210 AAA Lys	CAC His
	тўs	GIN	α	L225 AAA Lys	Arg	Leu	AAG Lys	TCC Ser	AGC Ser	Cys	Iys	AGA Arg	CAC	CCT Pro	Leu	Tyr	GTG Val	GAC Asp
40	1270 TTC Phe	AGT Ser	ASp	GIG Val	GGG	1285 TGG Trp	Asn	GAC Asp	TGG Trp	ATT	L300 GTG Val	GCT Ala	CCC Pro	ccc Pro	GGG	315 TAT Tyr	CAC His	GCC Ala
45	TTT Phe	TAC	330 TGC Cys	His	Gly	GAA Glu	TGC	345 CCT Pro	Phe	Pro	CTG Leu	GCT	.360 GAT Asp	CAT His	CTG . Leu .	AAC Asn	TCC	ACT Thr
50	AAT Asn	CAT His	GCC . Ala	ATT	.390 GIT Val	CAG Gln	ACG Thr	TTG (Leu '	GTC	AAC ASD	TCT Ser	GTT Val	AAC Asn	TCT .	420 AAG Lys	ATT Ile	CCT Pro	AAG Lys

1435 1450 1465 1480 GCA TGC TGT GTC CCG ACA GAA CTC AGT GCT ATC TCG ATG CTG TAC CTT GAC GAG Ala Cys Cys Val Pro Thr Glu Leu Ser Ala Ile Ser MET Leu Tyr Leu Asp Glu

1495 1510 1525 AAT GAA AAG GIT GIA TIA AAG AAC TAT CAG GAC ATG GIT GIG GAG GGI TGI GGG Asn Glu Lys Val Val Leu Lys Asn Tyr Gln Asp MET Val Val Glu Gly Cys Gly

1540(396) 1553 1563 1573 1583 1593 1603
TGT CGC TAGTACAGCA AAATTAAATA CATAAATATA TATATATATA TATATTTTTAG AAAAAAGAAA
Cys Arg

AAAA

and is of the following sequence:

5

10

15

25

40

45

50

55

Full-length hBMP-2Class II human cDNA clones are obtained in the following manner. The 200 bp EcoRI-SacI fragment from the 5' end of the Class II recombinant II-10-1 is isolated from its plasmid subclone, labeled by nick-translation, and hybridized to a set of duplicate nitrocellulose replicas of the U-2 OS cDNA library (25 filters/set; representing 500,000 recombinants). Hybridization and washing are performed under stringent conditions as described above. 16 duplicate positives are picked and replated for secondaries. Nitrocellulose filter replicas of the secondary plates are made and hybridized to an oligonucleotide which was synthesized to correspond to the sequence of II-10-1

CGGGCGCTCAGGATACTCAAGACCAGTGCTG

Hybridization is in standard hybridization buffer at 50°C with washing at 50° in 1 X SSC, 0.1% SDS. 14 recombinant bacteriophages which hybridize to this oligonucleotide are plaque purified. Plate stocks are made and small scale bacteriophage DNA preparations made. After subcloning 3 of these into M13, sequence analysis indicates that they represent clones which overlap the original Class II clone. One of these, lambda U2OS-3, was deposited with the ATCC under accession number 40342 on June 16, 1987. U2OS-3 contains an insert of approximately 1.8 kb. The partial DNA sequence and derived amino acid sequence of U2OS-3 are shown below in Table IV. This clone is expected to contain all of the nucleotide sequence necessary to encode the entire human BMP-2 Class II protein. This cDNA contains an open reading frame of 1224 bp, encoding a protein of 408 amino acids, preceded by a 5' untranslated region of 394 bp with stop codons in all frames, and contains a 3' untranslated region of 308 bp following the in-frame stop codon. The 8 bp region preceding the 5' untranslated region represents a linker used in the cDNA cloning procedure. This protein of 408 amino acids has molecular weight of 47kd and is contemplated to represent the primary translation product.

TABLE IV

							•
5	10 CTCTAGAGGG	20 CAGAGGAGGA	30 GGGAGGGAGG		50 GGAGCCCCGC	60 COGGAAGCTA	70 GGIGAGIGIG
10	80 GCATCCGAGC	90 TGAGGGACGC	100 GAGCCTGAGA	110 CCCCCTCCT	120 GCICCGGCIG	130 AGUATCIAGO	140 TIGICIOCOC
	150 GATGGGATTC	160 COSTOCAAGC	170 TATCTCGAGC	180 CIGCAGGGCC	190 ACAGTOCCOG	200 GCCCTCGCCC	210 AGGTTCACTG
15	220 CAACCGITCA	230 GAGGICCCCA	240 GGAGCTGCTG	250 CIGGOGAGOC	260 CGCTACTGCA	270 GGGACCTATG	280 GAGCCATTCC
20	290 GIAGIGOCAT	300 CCCCACCAAC	310 GCACTGCTGC	320 AGCITCOCIG	330 AGCCITTCCA	340 GCAAGITIGI	350 TCAAGATTGG
25	360 CIGICAAGAA	370 TCATGGACIG	ÓSÉ STATTATTATT		400 TGTCAAGACA	(1) CC ATG ATT MET Ile	
	417 GGT AAC CGA Gly Asn Arg	ATG CTG AT	432 G GIC GIT T T Val Val I	TIA TIA TGC Leu Leu Cys	447 CAA GIC CIG Gln Val Leu	46 CIA GGA GG Leu Glv Gl	C CCC
30	AGC CAT GCT Ser His Ala	477 AGT TIG AT	A TA COT GAG A	192 ACG GGG AAG	507 AAA AAA GIY	GOC GAG AT	T CAG
35	522 GGC CAC GCG Gly His Ala	53 GGA GGA CC	7 SC CGC TCA G	552 IGG CAG AGC	CAT GAG CTC	567 CTG CGG GA	C TIC
40	582 GAG GOG ACA Glu Ala Thr	CIT CIG CA	597 G ATG TIT G	og circ ogc	612 CCC CCC CCC	CAG CCT AG	627 C AAG
4 5	AGT GCC GTC Ser Ala Val	642 ATT CCG GA	C TAC ATG. C	657 CG GAT CIT	TAC OGG CIT	672 CAG TCT GG	G CAG
	687 GAG GAG GAA Glu Glu Glu	GAG CAG AT	702 C CAC AGC A	CT GGT CTT	717 GAG TAT CCT	73 GAG CGC CC	2 G GCC
50					-14 1/1 110	ozu my m	- 114

14

				747					762					777					
	AGC	ŒG	GCC			GTG:	ACC	. אכר			ראכ	מבט	CAA			CAG	220	ATC	
	Ser	Arq	Ala	Asn	Thr	Val	Arro	Ser	Dhe	His	His	Glu	וום ו	Hie	TAN	Gli	Aco	Ile	
								-	1110	*****	سيده		. GIG			GIU	, ve	116	
5	792					807					822					837			
	CCA	GGG	ACC	AGT	GAA			י ככידוי	بليليل	CCT			- वसम	ממ י	CTC	700) ACC	ATC	
	Pro	Gly	Thr	Ser	Glu	Asn	Ser	Ala	Phe	Arm	Phe	Teu	Pho	len	Ten	Sor	Cor	Ile	
		•															عص	116	
			852					867					882					897	
10	∞ T	GAG	AAC	GAG	GIG	ATC	TOO		GCA	GAG	CIT	œ			œ	GAG	CAG	GIG	
	Pro	Glu	Asn	Glu	Val	Ile	Ser	Ser	Ala	Glu	Leu	Arq	Leu	Phe	Arq	Glu	Gln	Val	
												_			_				
					912					927					942				
45	GAC	CAG	GGC	α	GAT	TGG	GAA	AGG	GGC	TTC	CAC	ŒI	ATA	AAC	ATT	TAT	GAG	GIT	
15	Asp	Gln	Gly	Pro	Asp	Trp	Glu	Arg	Gly	Phe	His	Arg	Ile	Asn	Ile	Tyr	Glu	Val	•
			•																
	ארוואר	957	~~~				972					987					1002		
	METT	AAG	<u> </u>	O'A	GCA	GAA	GIG	GIG	$\bar{\alpha}$	GGG	CAC	CIC	ATC	ACA	CCA	CIA	CIG	GAC	
20	PILI	тÃ2	PIO	PIO	ALA	GIU	val	Val	Pro	GIY	His	Leu	Ile	Thr	Arg	Leu	Leu	Asp	
			1	L017					1022										
	ACG	AGA			CAC	·cac	አአጥ	GIG	1032	~~	mcc.	CAR	. 3 cm	1047	C3.00	~~~	3.00		•
	Thr	Aru	Ten	Val	Hic	Hie	yen	Maj GTG	Thr	7~~	100	Clu	Mon	TIT	GAT.	GIG	AGC	CCI.	
						حسد	7311	Val	, <u></u>	ALG	u.	Gin	hin	me	ASP	.vai	ser	hto	
25	1062	2				1077				-	1092					1107			
	GCC	GTÇ	CTT	∞ c	TGG	ACC	ŒG	GAG	AAG.	CAG	CCA	·AAC	TAT	GGG	CTA	GCC	ייינים	GAG .	•
7	Ala	Val	Leu	Arg	Trp	Thr	Arg	Glu	Lys	Gln	Pro	Asn	Tyr	Gly	Leu	Ala	Ile	Glu	•
									_				-						
			122					1137				:	1152				3	1167	
30	37-1	ACT.	CAC	CIC	CAT	CAG	ACT	œ	ACC	CAC	CAG	GGC	CAG	CAT	GIC	AGG	ATT'	AGC	•
	val	TUE	HIS	Ieu	HIS	Gin	Thr	Arg	Thr	His	Gln	Gly	Gln	His	Val	Arg	Ile	Ser	
				1	182					107				_			٠		
	CGA	TCG	מיייני			CCC	y Cata	GGG	ת תחת ב	.197		C3.C	~m~	~~1	212	~~~			
	Ara	Ser	Len	Pro	Gln	Clv	Sor	Gly	yez.	TOG	712	CAR	Tou	7	<u>ccc</u>	CIC	CIG	GIC	
35						1	-	GLY	וובת	יביב	Ма	GTI	Leu	ALG	PIO	Teu	Leu	val	
	1	227				נ	242				1	257				,	272		
	∞	TTT	GGC	CAT	TAĐ	GGC	œ	GGC	CAT	GCC	TIG	ACC	CGA.	CCC	ന്ദ	AGG	GCC	AAC	
	Thr	Phe	Gly	His	Asp	Gly	Arg	Gly	His	Ala	Leu	Thr	Aru	Aru	Ara	Ara	Ala	Ivs	
40					_	_	_	_					3					,	
				287					302]	317					
	CT.	AGC	CCT	AAG	CAT	CAC	TCA	CAG	ŒG	GCC	AGG	AAG	AAG	AAT	AAG	AAC	TGC	ŒG	
	Arg	Ser	Pro	Lys	His	His	Ser	Gln .	Arg	Ala	Arg	Lys	Lys	Asn	Lys	Asn	Cys	Arg	
																		_	
45	1332		·			347					362				1	.377			
,	7	UAC Wic	Com	CIC	TAT'	GIG	GAC	TIC .	AGC	GAT	GIG	GGC	TGG	TAA	GAC	TGG	TTA	GIG	
	ALY .	حس	ser.	TEU.	ı AL	var	ASP	Phe	ser	ASP	val	GIĀ	JJ.	Asn	Asp	dtL	Ile	Val	
		ו	392				٦	407				,	422				_	457	
•	ccc			GGC (TAC	സ്ഥ		TIC	ጣልሶ	The state of the s	ርъጥ		.422 GNC	m-c	~~	CHAIN.	~~~	437	
50	Ala	Pro	Pro	Glv '	Tv~	Gln	λla	Phe '	1 472.	₩.	Wie Hie	Clv	yer.	760	D	III Dho	D	CIG	
			'	J	- 7			T 11C	4 Y L	UV3	حىدە	- V	A20	~v=	rit)	77 IE	rrt)	1 1-4 1	

	1452 1467 1482 GCT GAC CAC CTC AAC TCA ACC AAC CAT GCC ATT GTG CAG ACC CTG GTC AAT TCT Ala Asp His Leu Asn Ser Thr Asn His Ala Ile Val Gln Thr Leu Val Asn Ser	
5	1497 1512 1527 1542 GTC AAT TCC AGT ATC CCC AAA GCC TGT TGT GTG CCC ACT GAA CTG AGT GCC ATC Val Asn Ser Ser Ile Pro Lys Ala Cys Cys Val Pro Thr Glu Leu Ser Ala Ile	
10	1557 1572 1587 TCC ATG CIG TAC CIG GAT GAG TAT GAT AAG GIG GIA CIG AAA AAT TAT CAG GAG Ser MET Leu Tyr Leu Asp Glu Tyr Asp Lys Val Val Leu Lys Asn Tyr Gln Glu	
15	1602 1617 (408) 1636 1646 1656 ATG GTA GTA GAG GGA TGT GGG TGC CGC TGAGATCAGG CAGTCCTTGA GGATAGACAG MET Val Val Glu Gly Cys Gly Cys Arg	
	1666 1676 1686 1696 1706 1716 1726 ATATACACAC CACACACA CACCACATAC ACCACACACA	
20	1736 1746 1756 1766 1776 1786 1796 ACAGACIGCT TOCTTATAGC TEGACITTTA TITAAAAAAA AAAAAAAAAA AATEGAAAAA ATOCTTAAAC	
25	1806 1816 1826 1836 1846 1856 1866 ATTCACCTIG ACCITATIVA TGACTITACC TGCAAATGIT TTGACCATAT TGATCATATA TTTTGACAAA	
30	1876 1886 1896 1906 1916 1926 1936 ATATATTTAT AACTACGTAT TAAAAGAAAA AAATAAAATG AGTCATTATT TTAAAAAAAA AAAAAAACT	
	1946 CIAGAGINGA CEGAATIN	

CLAGAGIOGA OGGAATTC

The sequences of BMP-2 Class I and II as shown in Tables II, III IV and have significant homology to the beta (B) and beta (A) subunits of the inhibins. The inhibins are a family of hormones which are presently being investigated for use in contraception. See, A. J. Mason et al, Nature, 318:659-663 (1985). To a lesser extent they are also homologous to Mullerian inhibiting substance (MIS), a testicular glycoprotein that causes regression of the Mullerian duct during development of the male embryo and transforming growth factor-beta (TGF-b) which can inhibit or stimulate growth of cells or cause them to differentiate. Furthermore, the sequence of Table IV encoding hBMP-2 Class II has significant homology to the <u>Drosophila</u> decapentaplegic (DPP-C) locus transcript. See, J. Massague, Cell, 49:437-438 (1987); R. W. Padgett et al, Nature, 325:81-84 (1987); R.L. Cate et al, Cell 45: 685-698 (1986). It is considered possible therefore that BMP-2 Class II is the human homolog of the protein made from this transcript form this developmental mutant locus.

EXAMPLE VI

35

45

Expression of Bone Inductive Factors.

In order to produce bovine, human or other mammalian bone inductive factors, the DNA encoding it is transferred into an appropriate expression vector and introduced into mammalian cells by conventional genetic engineering techniques.

One skilled in the art can construct mammalian expression vectors by employing the sequence of Tables II, III AND IV or other modified sequences and known vectors, such as pCD [Okayama et al., Mol. Cell Biol., 2:161-170 (1982)] and pJL3, pJL4 [Gough et al., EMBO J., 4:645-653 (1985)]. The transformation of these vectors into appropriate host cells can result in expression of osteoinductive factors. One skilled in the art could manipulate the sequences of Tables II, III and IV by eliminating or replacing the mammalian regulatory sequences flanking the coding sequence with bacterial sequences to create bacterial vectors for intracellular or extracellular expression by bacterial cells. For example, the coding sequences could be further manipulated (e.g. ligated to other known linkers or modified by deleting non-coding sequences there-from or altering nucleotides therein by other known techniques). The modified bone inductive factor coding sequence could then be inserted into a known bacterial vector using procedures such as described in T. Taniguchi et al., Proc. Natl Acad. Sci. USA, 77:5230-5233 (1980). This exemplary bacterial vector could then be transformed into bacterial host cells and bone inductive factor expressed thereby. For a strategy for producing extracellular expression of bone inductive factor in bacterial cells., see, e.g. European patent application EPA 177,343.

Similar manipulations can be performed for the construction of an insect vector [see, e.g. procedures described in published European patent application 155,476] for expression in insect cells. A yeast vector could also be constructed employing yeast regulatory sequences for intracellular or extracellular expression of the factors of the present invention by yeast cells. [See, e.g., procedures described in published PCT application W086/00639 and European patent application EPA 123,289].

A method for producing high levels of an osteoinductive factor of the invention from mammalian cells involves the construction of cells containing multiple copies of the heterologous bone inductive factor gene. The heterologous gene can be linked to an amplifiable marker, e.g. the dihydrofolate reductase (DHFR) gene for which cells containing increased gene copies can be selected for propagation in increasing concentrations of methotrexate (MTX) according to the procedures of Kaufman and Sharp, <u>J. Mol. Biol.</u>, 159:601-629 (1982). This approach can be employed with a number of different cell types.

For example, a plasmid containing a DNA sequence for a bone inductive factor of the invention in operative association with other plasmid sequences enabling expression thereof and the DHFR expression plasmid pAdA26SV(A) 3 [Kaufman and Sharp, Mol. Cell. Biol., 2:1304 (1982)] can be co-introduced into DHFR-deficient CHO cells, DUKX-BII, by calcium phosphate coprecipitation and transfection. DHFR expressing transformants are selected for growth in alpha media with dialyzed fetal calf serum, and subsequently selected for amplification by growth in increasing concentrations of MTX (sequential steps in 0.02, 0.2, 1.0 and 5uM MTX) as described in Kaufman et al., Mol Cell Biol., 5: 1750 (1983). Transformants are cloned, and biologically active bone inductive factor expression is monitored by rat bone formation assay. Bone inductive factor expression should increase with increasing levels of MTX resistance. Similar procedures can be followed to produce other bone inductive factors.

Alternatively, the human gene is expressed directly, as described above. Active bone inductive factor may be produced in bacteria or yeast cells. However the presently preferred expression system for biologically active recombinant human bone inductive factor is stably transformed CHO cells.

30

35

50

55

As one specific example, to produce the human bone inductive factor (hBMP-1) of Example V, the insert of U2OS-1 is released from the vector arms by digestion with Sal I and subcloned into the mammalian expression vector pMT2CX digested with Xho I. Plasmid DNA from this subclone is transfected into COS cells by the DEAE-dextran procedure [Sompayrac and Danna PNAS 78:7575-7578 (1981); Luthman and Magnusson, Nucl. Acids Res. 11: 1295-1308 (1983)]. Serum-free 24 hr. conditioned medium is collected from the cells starting 40-70 hr. post-transfection.

The mammalian expression vector pMT2 Cla-Xho (pMT2 CX) is a derivative of p91023 (b) (Wong et al., Science 228:810-815, 1985) differing from the latter in that it contains the ampicillin resistance gene in place of the tetracycline resistance gene and further contains a Xhol site for insertion of cDNA clones. The functional elements of pMT2 Cla-Xho have been described (Kaufman, R.J., 1985, Proc. Natl. Acad. Sci. USA 82:689-693) and include the adenovirus VA genes, the SV40 origin of replication including the 72 bp enhancer, the adenovirus major late promoter including a 5' splice site and the majority of the adenovirus tripartite leader sequence present on adenovirus late mRNAs, a 3' splice acceptor site, a DHFR insert, the SV40 early polyadenylation site (SV40), and pBR322 sequences needed for propagation in E. coli.

Plasmid pMT2 Cla-Xho is obtained by EcoRl digestion of pMT2-VWF, which has been deposited with the American Type Culture Collection (ATCC), Rockville, MD (USA) under accession number ATCC 67122. EcoRl digestion excises the cDNA insert present in pMT2-VWF, yielding pMT2 in linear form which can be ligated and used to transform <u>E</u>. coli HB 101 or DH-5 to ampicillin resistance. Plasmid pMT2 DNA can be prepared by conventional methods. pMT2CX is then constructed by digesting pMT2 with Eco RV and Xbal, treating the digested DNA with Klenow fragment of DNA polymerase I, and ligating Cla linkers (NEBiolabs, CATCGATG). This removes bases 2266 to 2421 starting from the Hind III site near the SV40 origin of replication and enhancer sequences of pMT2. Plasmid DNA is then digested with EcoRl, blunted as above, and ligated to an EcoRl adapter,

5' PO4-AATTCCTCGAGAGCT 3'

3' GGAGCTCTCGA 5'

digested with Xhol, and ligated, yielding pMT2 Cla-Xho, which may then be used to transform <u>E. coli</u> to ampicillin resistance. Plasmid pMT2 Cla-Xho DNA may be prepared by conventional methods.

Example VII

Biological Activity of Expressed Bone Inductive Factor

5 A. BMP-1

To measure the biological activity of the expressed bone inductive factor. (hBMP-1) obtained in Example VI above. The factor is partially purified on a Heparin Sepharose column. 4 ml of transfection supernatant from one 100 mm dish is concentrated approximately 10 fold by ultrafiltration on a YM 10 membrane and then dialyzed against 20mM Tris, 0.15 M NaCl, pH 7.4 (starting buffer). This material is then applied to a 1.1 ml Heparin Sepharose column in starting buffer. Unbound proteins are removed by an 8 ml wash of starting buffer, and bound proteins, including BMP-1, are desorbed by a 3-4 ml wash of 20 mM Tris, 2.0 M NaCl, pH 7.4.

The proteins bound by the Heparin column are concentrated approximately 10-fold on a Centricon 10 and the salt reduced by diafiltration with 0.1% trifluoroacetic acid. The appropriate amount of this solution is mixed with 20 mg of rat matrix and then assayed for <u>in vivo</u> bone and cartilage formation as previously described in Example III. A mock transfection supernatant fractionation is used as a control.

The implants containing rat matrix to which specific amounts of human BMP-1 have been added are removed from rats after seven days and processed for histological evaluation. Representative sections from each implant are stained for the presence of new bone mineral with von Kossa and acid fuschin, and for the presence of cartilage-specific matrix formation using toluidine blue. The types of cells present within the section, as well as the extent to which these cells display phenotype are evaluated.

Addition of human BMP-1 to the matrix material resulted in formation of cartilage-like nodules at 7 days post implantation. The chondroblast-type cells were recognizable by shape and expression of metachromatic matrix. The amount of activity observed for human BMP-1 was dependent upon the amount of human BMP-1 protein added to the matrix. Table IX illustrates the dose-response relationship of human BMP-1 protein to the amount of bone induction observed.

Table IX

30	

35

IMPLANT N	IUMBER	AMOUNT USED (equivalent of ml transfection media)	HISTOLOGICAL SCORE
876-13	4-1	10 BMP-1	C+2
876-13	4-2	3 BMP-1	C+1
876-13	4-3	1 BMP-1	C+/-
876-13	4-4	10 MOCK	C-
876-13	4-5	3 МОСК	C-
876-13	4-6	1 MOCK	C -

Cartilage (c) activity was scored on a scale from 0(-) to 5.

Similar levels of activity are seen in the Heparin Sepharose fractionated COS cell extracts. Partial purification is accomplished in a similar manner as described above except that 6 M urea is included in all the buffers. Further, in a rat bone formation assay as described above, BMP-2 has similarly demonstrated chondrogenic activity.

The procedures described above may be employed to isolate other bone inductive factors of interest by utilizing the bovine bone inductive factors and/or human bone inductive factors as a probe source. Such other bone inductive factors may find similar utility in, inter alia, fracture repair.

The foregoing descriptions detail presently preferred embodiments of the present invention. Numerous modifications and variations in practice thereof are expected to occur to those skilled in the art upon consideration of these descriptions. Those modifications and variations are believed to be encompassed within the claims appended hereto.

50 Claims

Claims for the following Contracting States: BE, CH, DE, FR, GB, IT, LI, LU, NL, SE

1. A gene encoding human BMP-2 comprising the following DNA sequence:

		10			20		3	30		40)		50			60		70	
	GTOGACI	CIA	GAGI	GIGI	GT	CAGC	CITO	G C	TGGGG	ACTI	CTI	GAAC	TIG	CAGG	GAGA	AT A	AACTIX	3CGCA	
_					-							-							
5		-													-				
	~~~	80	~~	~~~	90	<del></del> ~	10			110			120		1	30		140	
	CCCCACI	TIL	u.u.	T-C-TC	<b>-</b> CC	11160		KG C	GGAGC	CIGO	TIC	GCC?	TCT	CCC2	reccc	XX (	$\infty$	CICC	
10		150		נ	60		17	70		180			190		-	00		210	
70	ACTOCIO	CGC	CITG			ACTGA			GTTCC			ΑΑΑΑ	CAG	AGAC	TICO:	מסי מרכי מ		210	
													~~		-1000	ω (	-	ancec.	
		220			30		24	0		250			260		2	70		280	
15	GGGAGAA	GGA	GGAG	GCAA	AG	AAAAC	GAAC	G G	ACATI	CCI	$\infty$ 1	TGCG	CCA	GGTC	CITI	GA (	CAGAC	TTTT	
								•						•		•	:	•	
	. <u>See a see a</u>	290,		3	ΩΩ.		27	Λisu:		220			222		٠. ـ	-			
	TCCATGI	GGA	CCT	CTTT	CA Z	מיייים מיייים	سي	TI A	···· .	320 320		רינונט ייייי	330	CNOT		40™ ™> π	and was	··350···	
						11001	IG	1 0	ccoc	GIGC	110	1 TWP	ج	GACT	عصع	IC 1	CCTAA	MGGT	
20		٠ ٠.				• • • •	·- ;• •									:	ur i e e		
		1)				370				3	85				4	00			
	OGACC A	TG G	TG G	$\infty$ G	GG 1	ACC C	GC T	GT (	CTT C	IA G	σт	TG C	TG C	TT C	$\propto c$	AG G	TC		
•	M	ET V	al A	la G	ly :	ihr A	rg C	ys 1	Leu L	eu A	la L	eu L	eu I	eu P	ro G	ln V	'al		
05			415						_										
25	CTC CTG	GGC			cm	r ccc	CTC	430	) "~~	C3.C	-m-		445						
	Leu Leu	Glv	Glv	Ala	Ala	Glv	Ten	Ul. Val	Dm	Clu	CIG	CIV	7	AGG	AAG	TTC	GCG		
		1	1		•	. 013	سجد	٧٣	rio	GIU	Leu	GIĀ	ALG	Arg	туs	me	AIA		
	460				475					490					505				
30	ccc ccc	TCG	TCC	GGC	$\alpha$	$\infty$	TCA	TO	CAG	$\infty$	TCT	GAC	GAG	GTC	CMG	AGC	GAG		
	Ala Ala	Ser	Ser	Gly	Arg	J Pro	Ser	Ser	Gln	$\mathbf{Pro}$	Ser	Asp	Glu	Val	Leu	Ser	Glu		
		<b>500</b>																	
	יייר כאכ	520	CCC.	cm-	CITY C		535			~~~		550					565		
	TTC GAG Phe Glu	Ten	Am	Tan	Ten	Sor	MET	The	GGC	CIG	AAA	CAG	AGA	$\infty$	ACC	$\infty$	AGC		
35	,		9		1,00	. Ser	PIET	File	GIY	LEU	гуs	GILI	Arg	Pro	ınr	Pro	Ser		
				580					595				•	610					
	AGG GAC		GIG	GTG	$\infty$	· ccc	TAC	ATG	CTA	GAC	CIG	TAT	œс	AGG	CAC	TCG	GGT		
	Arg Asp	Ala	Val	Val	Pro	Pro	Tyr	MET	Leu	Asp	Leu	Tyr	Arg	Arq	His	Ser	Glv		
										_		-	_	_			2		
40	625					640					655					670			
	CAG COG	GGC	TCA	$\infty$	GCC	CCA	GAC	CAC	: ccc	TTG	GAG	AGG	GCA	GCC	AGC	<b>Œ</b> A	CCC		
	Gln Pro	GIĀ	ser	PIO	Ата	oru	Asp	HLS	Arg	Leu	Glu	Arg	Ala	Ala	Ser	Arg	Ala		
			685					700	1				715						
45	AAC ACT	GIG		AGC	ىكىل	ר אַר	ርልጥ			יודאט	حلبت	CDA	715	מידי	~~	C	<b>3</b> ~		
73	Asn Thr	Val	Ara	Ser	Phe	His	His	Glu	Gin	Ser	Ten	COM	Chr	Len	Dm	Clu	Thr		
									. UIU		عانعم	GIU	GIU	LEU	TLU	GLU	TITE		

						745	,				760	,				775	`	
	AGI	GGG	AA.	ACZ	A ACC	: œc	AG	TT A	C TIX	C TT			AG	r TC	ר איני	$= \frac{1}{100}$	. ACC	GAG
5	Ser	Gl	/ Lys	Th	r Thr	Arg	Arg	g Phe	e Phe	Phe	e Asr	Leu	Ser	Se	r Ile	Pro	Thi	Glu
			790	)				805	5				820	)		-		835
	GAG	TTI	' ATC	) AC	TCA	GCA	GAC	cr:	CAC	GII	TIC	CGA	GAZ	CAC	S ATC	CAZ	A GAI	COT
	Glu	Phe	: Ile	Thr	: Ser	Ala	Gli	ı Let	ı Glr	ı Val	. Phe	Arg	Glu	ı Gli	n MEI	Glr	Asp	Ala
10																	_	
	ATT	GGA	אאר	רבב י	850 200		אחוו	י ראח	n ~~	865	) - 2011				880			
	Leu	Glv	Asn	Asr	AGC Ser	Ser	Phe	- CAI	CAC Hic	: Am	TIA	AAT	TIA	. TA	L GAA	YEA	ATA	AAA
		_							•	, ALG	, 116	, NSI	TTE	: TĂI	. GIU	TTE	; TTE	TAR
15		895					910					925					940	)
	CI	GCA	ACA	. c∞	AAC	TCG	AAA	TIC	$\infty$	GIG	ACC	AGI	CII	TIC	GAC	ACC	ACC	בתוחי
	Pro	Ala	Inr	· Ala	Asn	Ser	Lys	Phe	Pro	Val	Thr	Ser	Leu	Le	a Asp	Thr	Arg	Leu
				955					970									
20	GIG	AAT	CAG		GCA	AGC	AGG	: TCC	שונה ממט:	, yczr	المليان	ייי באיזי	CITY C	985	) . ~~	con	. ~	3000
20	Val	Asn	Gln	Asn	Ala	Ser	Aro	Tro	Glu	Ser	Phe	ASD	Val	ACC Thr	· Dm	ΔCI	. GIG	MEM
							-						٠		. 110	ALG	vai	. PÆI
	100					1015					1030					1045		
	222	TGG	ACT	GCA	CAG	GGA	CAC	. ecc	AAC	CAT	GGA	TTC	GIG	GIG	GAA	GIG	GCC	CAC
25	en tyeer i	110	шш	MIG	Gln	GIY	HIS	Ала	AST	His	GIY	Phe	Val	Val	Glu	Val	Ala	His
			1060			•		1075					1090					7705
	TIG	GAG	GAG	AAA	CAA	GGT	GTC	TOO	AAG	AGA	CAT	GTT	AGG	בידע	AGC	AGG	TY TH	1105
	Leu	Glu	Glu	Lys	Gln	Gly	Val	Ser	Lys	Arg	His	Val	Arg	Ile	Ser	Arg	Ser	Leu
30													_					
	CAC	<b>(2)</b>	CAT		1120	300	mcc	mos.	<b>~</b>	1135	3.00				1150	_		
	His	Gln	Asp	Glu	CAC His	Ser	TGG	Ser	Gla	ATA	AGG	CCA D	TIG	CLA	GIA	ACT	TTT	GGC
							111	اعاد	GILI	TIE	My	PLO	Leu	Leu	val	.mr	me	GIY
35		.165					180					1195					1210	
	CAT	GAT	GGA	AAA	GGG	CAT	$\alpha$	CTC	CAC	AAA	AGA	GAA	AAA	CCT	CAA	CCC	777	CAC
	HIS	Asp	Gly	Lys	Gly	His	Pro	Leu	His	Lys	Arg	Glu	Lys	Arg	Gln	Ala	Lys	His
				L225				_										
40	AAA	CAG			œc	CTT	AAG	TYY.	1240 260	ינבאנו	AAC	ስርስ	CAC	1255	mm-	ma o	~~~	~~
10	Lys	Gln	Arg	Lys	Arg	Leu	Lvs	Ser	Ser	CVS	Tus	ATTT	Hie	Dm	Ten	TAC	GIG	GAC
				-			_			1 _	-1-			110	Leu	TYL	Val	wah
	1270					285				3	300				נ	315		
	TTC .	AGI So~	GAC	GIG	GGG	TGG.	AAT	GAC	TGG	TTA	CIG	GCT	$\alpha$	$\infty$	GGG	TAT	CAC	GCC
45	Phe i	ser	Asp	vai	GIY.	Im.	Asn	Asp	Trp	Ile	Val	Ala	Pro	Pro	Gly	Tyr	His	Ala
		ו	330				ר	.345				,	260				_	
	TTT			CAC	GGA (	GAA '			ידידידי	CCT	CTG	പ്ര _ച	360 ლათ	ידב	CIIC	አአሮ	سب آ	.375
	Phe ?	Tyr ·	Cys	His	Gly	Glu	Cys	Pro	Phe	Pro	Leu	Ala .	Asp	His	Leu	Asn	Ser	Thr
50							-											
	አንጠ 4	73 CT			390				1	405				נ	420			
	AAT (	Al (	SCC A	AII' (	GIT (	AG I	ACG	TIG	GIC	AAC	TCT	GIT :	AAC	TCT	AAG .	TTA	CT.	AAG
	Asn I	ر حب	nud .	TTE	val (	י וווני	uIT	Leu	val.	ASN .	ser '	val 1	Asn .	Ser	Lys	Ile	Pro	Lys

1435 1450 1465 GCA TGC TGT GTC CCG ACA GAA CTC AGT GCT ATC TCG ATG CTG TAC CTT GAC GAG Ala Cys Cys Val Pro Thr Glu Leu Ser Ala Ile Ser MET Leu Tyr Leu Asp Glu 5 1510 AAT GAA AAG GIT GIA TIA AAG AAC TAT CAG GAC AIG GIT GIG GAG GGT TGT GGG Asn Glu Lys Val Val Leu Lys Asn Tyr Gln Asp MET Val Val Glu Gly Cys Gly 10 1573 1583 1593 TGT CGC TAGTACAGCA AAATTAAATA CATAAATATA TATATATATA TATATTITTAG AAAAAAGAAA Cys Arg AAAA 15 2. A gene encoding human BMP-2 having the amino acid sequence given in claim 1. 3. A gene encoding a protein exhibiting properties of human BMP-2 and comprising a DNA sequence: 20 (a) which differs from a DNA sequence of claim 1 in codon sequence due to the degeneracy of the genetic code; (b) which hybridises with a DNA sequence of claim 1 or section (a), above; or (c) represents a fragment, allelic or other variation of a DNA sequence of claim 1, whether said variation results in changes in the peptide sequence or not. 25 4. The DNA sequence of claim 3, which is a genomic DNA sequence. 5. The DNA sequence of claim 3, which is a cDNA sequence.

30 6. A gene encoding bovine BMP-2 comprising the following DNA sequence:

35

40

45

55

5	(1) GGC G	CAC H	GAT D	GGG G	15 AAA K	GGA G	CAC H	CCT P	CTC L	30 CAC H	AGA R	AGA R	GAA E	AAG K	45 CGG R
5		GCA A		CAC H		CAG Q			CGC R	75 CTC L	AAG K	TCC S	AGC S	TGT C	90 AAG K
10	AGA R	CAC H	CCT P	TTA L	105 TAT Y	GTG V	GAC D	TTC F	AGT S	120 GAT D		GGG G	TGG W	AAT N	135 GAC D
15	TGG W	ATC I	GTT V	GCA A	150 CCG P	CCG	GGG G	TAT Y	CAT H	165 GCC A	TTT F	TAC Y	TGC C	CAT H	180 GGG G
20	GAG E	TGC	CCT	TTT F	P	CTG L	GCC A	GAT D	CAC H	210 CTT L	AAC N	TCC S	ACG T	AAT N	225 CAT H
e est	GCC A	Ι	٧	Q	T	L	V	N	S	V	AAC N	S	'AAG K	I	P
25	AAG	GCA A	TGC	TGT	GTC	CCA P	ACA	GAG	CTC	AGC	GCC	ATC I	TCC	ATG	315 CTG L
30	TAC Y	CTT L	GAT D	GAG. E	330 AAT N	GAG E	AAG K	gtg V	GTA V	345 TTA L	AAG K			CAG O	
35	ATG M	GTT V	GTC V	GAG E	GGT	 TGT C	GGG	TGT	(129 CGT R	) TAGO	CACAC	97 GCA <i>F</i>	\AAT/	40 NAAAT	7 [A
	TAA	ATATA	117 ATA 1	rata:	42 TATAT	:7 :A TI	'AGA	437 \AAA(	AGC	<b>LAAA</b> S	447 AAA	TCAA	AGTTO	57 SAC	
40	ACT	, raati	167 TAT 1	TCCC	47 CAATG		ACTI	487 TATI			497 TGG	AATO		507 AA	
45	AAG	<b>LAAA</b> A	517. ACA C	AGCT	52 ATTT 57	T GA	AAAC				547 CTA	CCGA		557 SAA	
	GTT	GGAA		TAAA:			TCAG			TT					

50 7. A gene encoding bovine BMP-2 containing the amino acid sequence of claim 6.

· 55

- 8. A gene encoding a protein exhibiting properties of bovine BMP-2 and comprising DNA sequences:
  - (a) which differ from a DNA sequence of claim 7 in codon sequence due to the degeneracy of the genetic code;
  - (b) which hybridise with a DNA sequence of claim 7 or section (a), above; or
  - (c) represent fragments, allelic or other variations of a DNA sequence of claim 7, whether said variations result in changes in the peptide sequence or not.

- 9. The DNA sequence of claim 8, which is a genomic DNA sequence.
- 10. The DNA sequence of claim 8, which is a cDNA sequence.

5 11. A gene encoding human BMP-4 comprising the following DNA sequence:

	10			40			
	CTCTAGAGGG	CAGAGGAGGA	GGGAGGGAGG	GAAGGAGOGC	GGAGCCCGGC	COCCĂACCIA	GCIGAGIGIG
10							
	80	90		110		130	
	GCATCOGAGC	TGAGGGAGGC	GAGCCTGAGA	CCCCCIGCI	GCTCCCGCTG	AGTATCTAGC	TIGICICCC
15							
15					190		
	GATGGGATIC	COCTICCAAGC	TATCTOGAGC	CICCAGOGCC	ACAGTCCCCG	cocroscoc	AGGITCACIG
					260		
20	CAACCGTTCA	GAGGTOCCCA	GGAGCTGCTG	CIGGOGAGCC	OGCTACTGCA	GGGACCTATG	GAGCCATTCC
	_ 290				330		
	GIAGIGCCAT	CCCCACCAAC	GCACTGCTGC	ACCITOCCIG	AGCCTTTCCA	GCAAGITIGI	TCAAGATTGG
25		•					ing the second of the second o
	360	370	380	390	400	(1)	
	CTGTCAAGAA	TCATGGACIG	TEATTATATG	CCTTCTTTTC	TGTCAAGACA	CC ATG ATT	CCT
		:				MET .Ile	Pro

		417					432					447					462	
	GGT	AAC	<b>CGA</b>	ATG	CIG	ATG	GIC	GIT	TTA	TTA	TGC	CAA	GTC	CTG	CTA	GGA	GGC	GCG
	Gly	Asn	Arg	MET	Leu	MET	Val	Val	Leu	Leu	Cys	Gln	Val	Leu	Leu	Gly	Gly	Ala
5											_						_	
				477					492					507				
	AGC	CAT	GCT	AGT	TIG	ATA	$\infty$ T	GAG	ACG	GGG	AAG	AAA	AAA	GTC	GCC	GAG	TTA	CAG
	Ser	His	Ala	Ser	Leu	Ile	Pro	Glu	Thr	Gly	Lys	Lys	Lys	Val	Ala	Glu	Ile	Gln
	Enn																	
10	522	CT C		~~1	~~	537	~~~				552					567		
	Cly	LAC	712	Clar	Clar	2000	3	TCA	GGG	CAG	AGC	CAT	GAG	CIC	CIG	₩.	GAC	TTC
	Gly	шѕ	Ala	GIY	GTA	Arg	Arg	Ser	GIA	Gin	Ser	HIS	GIU	Leu	Leu	Arg	Asp	Phe
			582					597					612					627
	GAG	GŒ		CIT	CTG	CAG	ATG		GGG	cmc	CCC	CCC	CCC	œ	CNG	<del>Д</del>	NCC	627 AAG
15	Glu	Ala	Thr	Leu	Leu	Gln	MET	Phe	Glv	Teu	Am	Am	Am	Pm	CAB	Dm	Sor	Lys
									<b></b> 1		9	9	1119	110	GHI	PLO	561	Lys
					642					657					672			
	AGT	GCC	GTC	ATT	$\infty$	GAC	TAC	ATG	CCC	GAT	CTT	TAC	œ	CTT	CAG	тст	GGG	GAG
	Ser	Ala	Val	Ile	Pro	Asp	Tyr	MET	Arg	Asp	Leu	Tyr	Arq	Leu	Gln	Ser	Glv	Glu
20						_	_		_	-		•	•					
		687					702					717					732	
	GAG	GAG	GAA	GAG	CAG	ATC	CAC	AGC	ACT	CGI	CTT	GAG	TAT	CCT	GAG	$\alpha$ C	$\infty$	GCC
	Glu	Glu	Glu	Glu	Gln	Ile	His	Ser	Thr	Gly	Leu	Glu	Tyr	Pro	Glu	Arg	Pro	Ala
										•								
25				747					762					777				
	AGC	œ	GCC	AAC	ACC	GIG	AGG	AGC			CAC	GAA	GAA	CAT	CTG	GAG	אאר	ATC
	Ser	Arg	Ala	Asn	Thr	Val	Arg	Ser	Phe	His	His	Glu	Glu	His	-Leu	-Glu	Asn	Ile
							_											
20	792					807					822					837		
30	CCA	GGG	ACC	AGT	GAA	AAC	TCT	GCT	TTT	ŒT	TTC	CTC	TIT	AAC	CIC	AGC	AGC	ATC
	Pro	GTĀ	Thr	Ser	Glu	Asn	Ser	Ala	Phe	Arg	Phe	Leu	Phe	Asn	Leu	Ser	Ser	Ile
			050															
	ىلىک	GNG	852	CAC	cmc	3000	m~	867		~~	~		882					897
35	Pm	Glu	yen	Chi	Lett.	TIO	200	TCT	SCA	CAG	CIT	CGG	CIC	TIC	œ	GAG	CAG	GIG
			7-341	Giu	٧	TTE	Ser	Ser	Ma	GIU	Leu	Arg	Leu	me	Arg	GIU	GIn	Val
					912					927					043			
	GAC	CAG	GGC	CT		TGG	GAA	AGG	GGC	TTC	CAC	CT	ልጥል	220	942	ייים	GNG	GIT
	Asp	Gln	Gly	Pro	Asp	Trp	Glu	Arq	Gly	Phe	His	Arg	Tle	Asn	Tle	Tyr	Glu	Val
40			_		_	•		_	-			5				-1-	O_L	vu
		957					972					987				-	1002	
	ATG	AAG	$\infty$	$\infty$ A	GCA	GAA	GTG	GTG	$\alpha$	GGG	CAC	CTC	ATC	ACA	<b>CGA</b>	CTA	സ്ത	GAC
	MET	Lys	Pro	Pro	Ala	Glu	Val	Val	Pro	Gly	His	Leu	Ile	Thr	Arq	Leu	Leu	ASD
																		•
45	300	3.63		1017				]	L032				_ ]	047				
	AU	ALA	CIG	GIC	CAC	CAC	AAT	GIG	ACA	ŒG	TGG	GAA	ACT	TTT	GAT	GIG	AGC	$\alpha$
٠'	ши	Arg	Leu	vai	HIS	HIS	Asn.	Val	Thr	Arg	Trp	Glu	Thr	Phe	Asp_	Val	Ser	Pro
	1062							•							_		,	
			ىئىلى	CCC		.077	~~	CNC	220	C2 ~	092	220	<b></b>		1	107		
50 ':	GCG	Vai	Len	کیت Am	J.M.	حديل	ميد	Clu.	rickt. Tive	CJ~	D-M	AAC Acm	TAT.	GGG	CLA.	GCC.,	ATT.	GAG
					-~P		9	JIU	בעב	3111	FIO	ווכה	TĀL	GTĀ	⊥eu.	мта	тте	GIU
		1	122				1	.137				ר	152				,	167
	GIG .			CIC	CAT	CAG	ACT	ŒG	ACC	CAC	CAG	GGC	CAG	САТ	CITC	እርር	νωυ. Τ	167 300
	Val	Thr	His	Leu	His	Gln	Thr	Arq	Thr	His	Gln	Glv	Gln	His	Val	Am	Tle	Ser
ee.								_				_				-7		

	1182 1197 1212
_	OGA TOG TTA OCT CAA GGG AGT GGG AAT TGG GGC CAG CTC GGG CCC CTC CTG GTC
5	Arg Ser Leu Pro Gln Gly Ser Gly Asn Trp Ala Gln Leu Arg Pro Leu Leu Val
	1227 1242 1257 1272
	ACC TIT GGC CAT GAT GGC GGG GGC CAT GCC TIG ACC GGA GGC GCC ACC
	Thr Phe Gly His Asp Gly Arg Gly His Ala Leu Thr Arg Arg Arg Arg Ala Lys
10	3307
	OGT AGC CCT AGC CAC TCA CAG CGG GCC AGG AAG AAG AAT AAG AAC TGC CGG
	Arg Ser Pro Lys His His Ser Gln Arg Ala Arg Lys Lys Asn Lys Asn Cys Arg
	****
15	1332 1347 1362 1377
	CGC CAC TOG CTC TAT GTG GAC TTC AGC GAT GTG GGC TGG AAT GAC TGG ATT GTG Arg His Ser Leu Tyr Val Asp Phe Ser Asp Val Gly Trp Asn Asp Trp Ile Val
	3 and and and the set with Asia aspared the set
	1392 1407 1422 1437
	GCC CCA CCA GGC TAC CAG GCC TTC TAC TGC CAT GGG GAC TGC CCC TTT CCA CTG
20	Ala Pro Pro Gly Tyr Gln Ala Phe Tyr Cys His Gly Asp Cys Pro Phe Pro Leu
	1452 1467 1482
	GCT GAC CAC CTC AAC TCA ACC AAC CAT GCC ATT GTG CAG ACC CTG GTC AAT TCT
	Ala Asp His Leu Asn Ser Thr Asn His Ala Ile Val Gln Thr Leu Val Asn Ser
25	1497 1512 1527 1542
	GTC AAT TOO AGT ATC OOC AAA GOO TGT TGT GTG COO ACT GAA CTG AGT GOO ATC
	Val Asn Ser Ser Ile Pro Lys Ala Cys Cys Val Pro Thr Glu Leu Ser Ala Ile
	3572
30	1557 1572 1587 TCC ATG CTG TAC CTG GAT GAG TAT GAT AAG GTG GTA CTG AAA AAT TAT CAG GAG
	Ser MET Leu Tyr Leu Asp Glu Tyr Asp Lys Val Val Leu Lys Asn Tyr Gln Glu
	1602 1617 (408) 1636 1646 1656
	ATG GTA GTA GAG GGA TGT GGG TGC CGC TGAGATCAGG CAGTCCTTGA GGATAGACAG MET Val Val Glu Gly Cys Gly Cys Arg
35	The var var old dry cys dry cys Arg
	1666 1676 1686 1696 1706 1716 1726
	ATATACACAC CACACACACA CACCACATAC ACCACACA CACCITCOCA TOCACTOACO CACACACTAC
40	1736 1746 1756 1766 1776 1786 1796
	ACAGACIGCT TOCTTATAGC TGGACTITTA TITTAAAAAA AAAAAAAAA AATGGAAAAA ATCCCTAAAC
	and the electric term of the common angles where a constitution by a part of the constitution of the const
	1006
45	1806 1816 1826 1836 1846 1856 1866 ATTCACCITG ACCITATINA TGACITIACO TGCANATGIT TTGACCATAT TGATCATATA TTTTGACANA
	TOTAL
	·
	1876 1886 1896 1906 1916 1926 1936
	ATATATTIAT AACTACGIAT TAAAAGAAAA AAATAAAATG AGTCATTATT TTAAAAAAAA AAAAAAAACT
50	
	1946
	CIAGAGICGA CGGAATIC
55	
	12. A gene encoding human BMP-4 having the amino acid sequence given in claim 11

^{12.} A gene encoding numan BMP-4 having the amino acid sequence given in claim 11.

^{13.} A gene encoding a protein exhibiting properties of BMP-4 and comprising a DNA sequence:

- (a) which differs from a DNA sequence of claim 11 in codon sequence due to the degeneracy of the genetic code;
- (b) which hybridises with a DNA sequence of claim 11 or section (a), above; or
- (c) represents a fragment, allelic or other variation of a DNA sequence of claim 11, whether said variation results in changes in the peptide sequence or not.
- 14. The DNA sequence of claim 13, which is a genomic DNA sequence.
- 15. The DNA sequence of claim 13, which is a cDNA sequence.
- 16. A vector containing the gene or DNA sequence of any one of claims 1 to 15 in operative association with an expression control sequence.
- 17. A cell transformed with a vector of claim 16.
- 18. The cell of claim 17 which is a mammalian cell, a bacterial cell, an insect cell or a yeast cell.
- 19. The cell of claim 18 which is a CHO cell.
- 20. A protein exhibiting properties of BMP-2 which is encoded by a gene or DNA sequence of any one of claims 1 to 10.
  - 21. A protein exhibiting properties of BMP-2, which is obtainable by the steps of culturing in a suitable culture medium a cell transformed with an expression vector comprising a gene or a DNA sequence of any one of claims 1 to 10, and recovering said protein from said culture medium.
  - 22. A protein exhibiting properties of BMP-4 which is encoded by a gene or DNA sequence of any one of claims 11 to 15.
  - 23. A protein exhibiting properties of BMP-4, which is obtainable by the steps of culturing in a suitable culture medium a cell transformed with an expression vector comprising a gene or a DNA sequence of any one of claims 11 to 15, and recovering said protein from said culture medium.
  - 24. A process for producing the protein of claims 21 or 23, comprising the steps of culturing in a suitable culture medium the cell of claim 17 and isolating said protein from said culture medium.
- 35 25. A pharmaceutical composition comprising the proteins of any one of claims 20 to 23, individually or in combination, and a pharmaceutically acceptable vehicle.
  - 26. The pharmaceutical composition of claim 25, further comprising a matrix capable of delivering the composition to the site of the bone or cartilage defect and providing a structure for inducing bone or cartilage formation.
  - 27. The pharmaceutical composition of claim 26, wherein said matrix comprises hydroxyapatite, collagen, polylactic acid or tricalcium phosphate.
- **28.** Use of a protein of any one of claims 20 to 23, individually or in combination, for the preparation of a pharmaceutical composition for inducing bone or cartilage formation.

#### Claims for the following Contracting State: AT

A process for the preparation of a gene encoding human BMP-2 comprising the following DNA sequence:

5

10

15

25

30

_	GIC	GACI	10 CIA	GAGI	GIGI	20 GT (	CAGCA		O CCI	CCCC	40 ACII		GAAC	50 TTG	CAGG		60 AT <i>1</i>	ACTI	70 303CA
5	<u></u>	CACI	80 TTG	œœ		90 CC 1	MIGC	10 XXXX		GAGC	CIGO 110			120 TCT	CCGA		30 CA (		140 CTC
10	ACT		150 GGC	CTTG		.60 AC #	ACTGA	17 GAGG		TTC	180 XAGC			190 GAG	AGAC		00 05 (	5005G	210 CACCC
15	GGG		220 GGA	GGAG		30 AG <i>P</i>	<b>LAAA</b> G	24 GAAO		CATT	250 CGT			260 CCA	GGTC	2 CITD	70 [/] GA (	CAGA	280 GITIT
	TCC		29.0 GGA	CCT		00. CA A	atoca Aesota	31 OGIG			320 320			330 AŒ	GACT		40 IC I	CCIA	350 AAGGT
20																			
	OGA	$\infty$ à	1) TG G ET V	TG G	∝ G la G	GG A	70 CC C hr A	rd ci ec 11	gr c ys L	TT C eu L	TA G	85 S T la L	TG C	IG C eu I	TT O eu P	$\propto \dot{a}$	oo AG C Ln V	TC al	
25				415					430					445		•			
	CIC	CIG	GC	GC	GCG Ala	GCT Ala	Gly	CIC	GTT Val	Pro	GAG Glu	CTG Leu	GLY	Arg	Arg	aag Lys	TTC	CCG Ala	
	460					475					490					505			
30																		GAG Glu	
			520					535					550					565	
	TIC	GAG	TIG	œ	CIG	CTC	AGC	ATG	TTC	GGC	CIG	AAA	CAG	AGA	$\infty$	ACC	$\infty$	AGC	
35	Pne	Glu	Leu	Arg	Leu	Leu	Ser	MET	Phe	Gly	Leu	Lys	Gln	Arg	Pro	Thr	Pro	Ser	
					580					595					610				
																		GIY	
	,			****	•	110	110	-1-	1		LOD		*3*	'ny	πg	ıus	الله	GIJ	
40	~ ~	625					640					655					670		
																		GCC	
			•										- 3				3		
	220	λCT	ست	685	200	~~~	03.0	C>T	700	C>>	W		C	715		~~~	CAR	<b>N</b> CC	
45																		ACG Thr	
				- 3							_								

	730 AGI	_GG	S AA	A AC	A .AC	74: C CC	G AG	A TT	e TTV	c TT	760 T-AA	מודוף יו	a.c.ı	ר-ער-יו	ייי אייי	775	3~	GAG
5	Ser	Gly	L	2 III	r Th	r Ar	Arg	g Phe	e Pho	e Ph	e Ası	n Leu	Ser	Se	r Ile	Pro	Thr	Glu
	GAG	TTT	79. 'AT	C AO	C TC	A GC	GA(	805 CIT	ר כאנ	GI	r TIX	C CGP	820 GAA	CNC	ota e	CAA	GAT	835 GCT
10	GIU	File	= 111	e mi	se 85	r Ala	i GII	ı Lei	ı Glr	n Val	l Phe	Arg	(Glu	(Glr	ı MEI	Gln	Asp	Ala
	TTA Leu	GIY	AA A	AA C ASI	r ag	C AG	TTC Phe	CAT His	CAC His	865 COG Arg	יייע ג	AAI AST	ATI Ile	TAI Tyr	880 GA2 Glu	3.000	ATA Ile	AAA Lys
15	CCT Pro	895 GCA Ala	AC	A GO	C AA A As	C TOO	910 AAA Lys	TTC	e exc	GIC Val	ACC Thr	925 AGI Ser	<u> </u>	TIC	GAC Asp	ACC Thr	940 AGG Arg	TIG Leu
20	GIG	AAT	CAC	955 AA1	5 1' GC2	A AGO	AGG	TGG	970 CAA	) ACT	ىلىلىك ر	G D TT	CITY:	985	. ~~	com	CTTC	ATG MET
	1000	)				1015					1030				H=	1045		
25	Arg	Trp	Thr	Ala	Glr	Gly	His	Ala	Asn	His	GLA	Phe	Val	Val	GAA Glu	GIG Val	GCC Ala	CAC His
	TTG Leu	CAG	1060 GAG Glu	AAA	CAA	GGI	GTC	1075 TCC	AAG	AGA	CAT	CITY	1090 AGG	ATA	AGC	AGG	mm.	1105 TTG
30					1120		•	<b>U</b> C <u>L</u>		بير 1135		var	Arg		ser 1150	Arg.	Ser	Ten
	CAC His	CAA Gln	GAT Asp	GAA	CAC	AGC	TGG Trp	TCA Ser	CAG	ATA	AGG	CCA Pro	TTG Leu	CTTS	עיוד	ACT Thr	TTT Phe	GC Gly
35	CAT	165 GAT	GGA	AAA	GGG	CAT	180 CT	crc	CAC	AÁA	)C)	195 GAA	222	رت	CD I	-	210	CAC.
	His .	Asp	GTÄ	тĀг	Gly	His	Pro	Leu	His	Lys	Arg	Glü	Lys	Arg	Gln	Ala	Lys	His
40	AAA (	CAG Gln	$\infty$ G	L225 AAA Lys	CGC Arg	CTT Leu	aag Lys	TOC	240 AGC Ser	TGT Cys	AAG Lys	AGA Arg	CAC	255 CCT Pro	TTG Leu	TAC Tyr	GIG ( Val :	GAC Asp
	1270 TTC 1	GT -	GAC	ണ്ട		1285 TGG	<u>አ</u> ልጥ	GàC I	TTC:C	1 2000	300	ccm	<b>~</b>	~~	1	315		
45	Phe S	Ser .	Asp	Val	Gly	Trp	Asn .	yzb ,	Trp	Ile	Val .	Ala	Pro :	Pro	Gly	TAT (	CAC ( His 1	Ala
50	TTT 1 Phe T	AC !	Cys IGC 330	CAC His	GGA Gly	GAA ' Glu '	IGC (	345 CCT ( Pro 1	TTT Phe	CCT Pro	CTG ( Leu )	GCT (	360 GAT ( Asp 1	CAT	CIG . Leu .	AAC : Asn :	m 2	375 ACT Thr
	AAT C Asn H	AT ( is )	SCC .	TTA	390 GTT Val	CAG I	ACG : Thr 1	MG ( Leu 1	FIC I	405 AAC : Asn :	ICT ( Ser V	FTT / /al /	AAC 1 Asn S	וכדי	420 AAG 1 Lys :	ATT ( Ile I	XT &	AG Lys

1435 1450 1465 1480 GCA TGC TGT-GTC CCG-ACA-GAA-CTC-AGT-GCT ATC TCG ATG-CTG TAC CTT GAC GAG Ala Cys Cys Val Pro Thr Glu Leu Ser Ala Ile Ser MET Leu Tyr Leu Asp Glu 5 1510 1525 1495 ANT CAA AAG GIT GTA TTA AAG AAC TAT CAG GAC ATG GIT GIG GAG GGT TGT GGG Asn Glu Lys Val Val Leu Lys Asn Tyr Gln Asp MET Val Val Glu Gly Cys Gly 10 1540 (396) 1553 1563 1573 1583 1593 TGT CGC TAGTACAGCA AAATTAAATA CATAAATATA TATATATATA TATATTTTAG AAAAAAGAAA Cys Arg

15 AAAA ,

20

30

40

45

50

55

wherein said process comprises the following steps:

- a) screening of a gene library constructed from U-2 OS derived DNA or cDNA with a labelled bBMP-2 fragment by hybridization,
- b) isolating positive clones, and
- c) isolating the DNA-inserts from said clones.
- 2. The process according to claim 1, wherein the gene encodes human BMP-2 having the amino acid sequence given in claim 1.
  - 3. A process for the preparation of a gene encoding a protein exhibiting properties of human BMP-2 and comprising a DNA sequence:
    - a) which differs from a DNA sequence of claim 1 in codon sequence due to the degeneracy of the genetic code;
    - b) which hybridizes with a DNA sequence of claim 1 or section (a), above; or
    - c) represents a fragment, allelic or other variation of a DNA sequence of claim 1, whether said variation results in changes in the peptide sequence or not,
- 35 wherein said process comprises standard techniques of molecular biology.
  - 4. The process according to claim 3, wherein the DNA sequence is a genomic DNA sequence.
  - 5. The process according to claim 3, wherein the DNA sequence is a cDNA sequence.
  - 6. A process for the preparation of a gene encoding bovine BMP-2 comprising the following DNA sequence:

5	(1) GGC G	CAC H	GAT D	GGG G	15 AAA K-	GGA G	CAC H	CCT P	CTC L	30 CAC H	AGA R			AAG K	
	CAA Q	GCA A	aaa K	CAC H	60 <b>AAA</b> K				CGC R	75 CTC L	AAG K	TCC S	AGC S	TGT C	90 AAG K
10	AGA R	CAC H	CCT P	TTA L	105 TAT Y	gtg V	GAC D	TTC F	agt S	120 GAT D		GGG G	TGG W	AAT N	135 GAC D
15	TGG W	ATC I	GTT V	GCA A	150 CCG P	CCG P	GGG G	TAT Y	CAT H	165 GCC A	TTT	TAC		CAT H	
20	GAG E	TGC C	CCT P	TTT	P	CTG L	GCC A	GAT D	CAC	210 CTT L	AAC N	TCC S		AAT N	225 CAT H
25	GCC A	ATT I	CTC V	CAA	240 ACT	CTG:	GTC V	AAC N	TCA S	255 GTT V	AAC	TCT S	AAG K		270 CCC P
	AAG K	GCA A	TGC C	TGT C	385 GTC V	CCA P	ACA T	GAG E	CTC L	300 AGC S	GCC A				315 CTG L
30	TAC Y	CTT L		GAG E		GAG E	AAG K	GTG V	GTA V	345 TTA L	AAG K			CAG O	
35	ATG M	GTT V	GTC V	GAG E	375 GGT <u>G</u>	TGT	GGG	TGT	(129 CGT R	TAGO	3 ACAG	97 CA A	AATA	4C AAAT	7 'A
40	TAAA	TATA		'ATAT		A TT		AAAC	AGC	AAAA	AAA	TCAA	GTTG		
•	ACTT		67 AT T	TCCC	47 AATG					GGAA	497 .TGG		5 GAGA		
45	AAGA		17 CA C	AGCT	52 ATTT		AAAC	537 TATA		'ATAT	547 CTA			57 AA	
	GTTG	_	67 AA C	AAAT.	57 ATTT		TCAG	587 AGAA		TT,					

wherein said process comprises the following steps:

a) screening a gene library constructed from bovine liver DNA or cDNA with a labelled probe designed on the basis of the amino acid sequence of a fragment of bBMP-2,

b) isolating positive clones, and

c) isolating the DNA-inserts from said clones.

^{7.} The process according to claim 6, wherein the gene encodes bovine BMP-2 having the amino acid sequence of claim 6.

	8.	A process for the preparation of a gene encoding a protein exhibiting properties of bovin DNA sequences:	e BMP-2 and comprising
5	-	<ul> <li>a) which differ from a DNA sequence of claim 7 in codon sequence due to the degene</li> <li>b) which hybridize with a DNA sequence of claim 7 or section a), above; or</li> <li>c) represent fragments, allelic or other variations of a DNA sequence of claim 7, when in changes in the peptide sequence or not,</li> </ul>	
10		wherein said process comprises standard techniques of molecular biology.	
10	9.	The process according to claim 8, wherein the DNA sequence is a genomic DNA seque	nce.
	10.	The process according to claim 8, wherein the DNA sequence is a cDNA sequence.	
15	11.	A process for the preparation of a gene encoding human BMP-4 comprising the following	g DNA sequence:
00			
20			
			•
25			
			•
			·
30			•
			•
35			
40			
40			
45			
50			
50			

		10			20		•	30			40		50			60		70
	CICIA	GAGGC	CA	2242	AGGA	GGG	AGGG	AGG	CAAC	GAGO	GC G	GAGO	XX	; ça	GĀA	CIA	GGI	AGIGIG
5																		
•		80			90			100			10		120			130		140
	GCATO	CGAGC	TG	AGGG	ACCC	GAG	CIG	AGA	$\infty$	CIG	CT G	CICC	GCTC	AG	[ATC]	AGC	TIG	CICCC
10		150			160			170		_	30		190			200		210
,,	GAIGG	GATIC	: œ	FICC.	AAGC	TAT	CICC	AGC	CIGC	AGCG	X A	CAGIY		GC	xia.	æ	AGG.	TCACTG
	~~~	220		~~~~	230			240	~~~		50		260			270	~2 ^r	280
15	CAACU	SITC	L GAI	3G1C	JUA	GGA	SCIG	CIG	CIEC	ZEAGK	.c 0	GCTA	CIGCA	GGC	ACCI	AIG	GALK	CATTCC
	CIBCI	290 זימיירי≎		~3~	300			310			20 N 3	~~~	330		الله الم	340		350 AGATTGG
	GINGI	بمب			·	بمب	-16-	100	MISCI	سب	·	GCC1.	LICCA	ر بحت	MEST 1	1761	ica	nivi 100
20								<u> </u>					بد ند ^م د		4-			
	CIGIC	360 44044		ATTCC	370°			380			90 സെന				(1) ATG	ינידע	CCT!	
						2 4.1		40									Pro.	
															•			
25																		
	com	417 AAC	~~ x	3.000	- TIV-	אתר	432	ست	עידיא	TTES	TCC	447	arc.	സ	בידים	സ്മ	462 GGC	സ
	Gly	Asn	YZG.	MET	Leu	MET	Val	Val	Leu	Leu	Cys	Gln	Val	Leu	Leu	Gly	Gly	Ala
	•						•				_					•		
30	ACC	CAT	CCTP	477	TUTE	בידע	رت.	CAC	492	GGG	AAG	AAA	AAA	507 GTC	GCC	GAG	ATT	CAG
	Ser	His	Ala	Ser	Leu	Πe	Pro	Glu	Thr	Gly	Lys	Lys	Lys	Val	Ala	Glu	Ile	Gln
											552					567		
	522 GGC	CAC	GCG	GGA	GGA	537 CGC	CGC.	TCA	GGG	CAG	AGC	CAT	GAG	CTC	CIG	œ	GAC	TTC
35	Gly	His	Ala	Gly	Gly	Arg	Arg	Ser	Gly	Gln	Ser	His	Glu	Leu	Leu	Arg	Asp	Phe
			582					597					612					627
	GAG	ထေ	ACA	CIT	CIG	CAG	ATG	TIT	GGG	CIG	œc	œc	ŒC	∞	CAG	α r	AGC	AAG
	-Glu	Ala	Thr	Leu	Leu	Gln	MET	Phe	GJĀ	Leu	Arg	Arg	yrd	Pro	Gln	Pro	Ser	Lys
40					642					657					672			
	AGT	ccc	GIC	ATT	œ	GAC	TAC	ATG	œ	GAT	CIT	TAC	ŒG	CIT	CAG	TCT	GGG	GAG
	Ser	Ala	Val	Ile	Pro	Asp	Tyr	MET	Arg	Asp	Leu	Tyr	Arg	Leu	Gln	Ser	Gly	Glu
		687					702					717					732	
45	GAG	GAG	GAA	GAG	CAG	ATC	CAC	AGC	ACT	GGT	CIT	GAG	TAT	CCT	GAG	œc	∞	ccc
	Glu	Glu	Glu	Glu	Gln	Ile	His	Ser	Thr	Gly	Leu	Glu	TYT	Pro	Glu	Arg	Pro	ALA

				747					762					777				
																	AAC Asn	
5	792					807					822					837		
											TTC					AGC	AGC Ser	
	110	CLI		Ser	Gia	Y21	Ser			ALG	FILE	LEU		יובא	LEG	54		
10																	CAG	
	Pro	Glu	ASN	Glu	Val	Ile	Ser	Ser	Ala	Glu	Leu	Arg	Leu	Phe	Arg	Glu	Gln	Val
	GAC	CAG.	ccc	Cut.	912 CAT	സാം	CAA) CC	ccc	927	CA C	CTI.	מינימ	220	942	ጥልጥ	GAG	مبيت
15																	Glu	
		957					972					987					1002	
																	CTG	GAC Asp
20				L017			-									-		
	ACG	AGA			CAC	CAC	AAT		LO32 ACA	œ	TGG	CAA		TTT	GAT	GIG	AGC	CCT
	Thr	Arg.	Leu	Val	His	His	yziu	.Val	Thr	Arg,	طنث	Glu	Thr	Phe	Asp	Val	Ser	Pro
25	1062					L077	·				092					1107		
																		GAG Glu
		נ	122				ב	137				נ	152]	1167
30		ACT	CAC				ACT	∞ G				GGC	CAG				ATT	AGC
	vai	шп	ur2	TEII	ris	σIJ	ımı	Arg	ш	ALS	GLN	GīĀ	GLII	nis	vai	Arg	Ile	ser.
				7	182				,	197					1212			
35	CGA	TCG	TTA	α r	CAA	œ	AGT	GGG	TAA	TGG	œ	CAG	CTC	œ	∞	CIC	CIG	GTC
	Arg	Ser	Leu	Pro	Gln	Gly	Ser	Gly	Asn	Trp	Ala	Gln	Leu	Arg	Pro	Leu	Leu	Val
		227	ccc	Carr	CAT		242	ccc	ርልጥ	ccc		1257	œ»	œc	œ	-	1272 GCC	AAG
40																	Ala	
				287					L302					1317				
	-CGT Arg	AGC Ser	CCT	AAG Lvs	CAT	CAC His	TCA Ser	CAG	CGG Arg	GCC Ala	AGG Arg	AAG Lvs	AAG Lvs	AAT Asn	AAG Lys	AAC Asn	TGC Cys	Arg
45	1332			-,-					5		1362					1377		_
	α C	CAC	TCG	crc	TAT	347 GTG	GAC	TTC	AGC	GAT	GTG	GGC	TGG	AAT	GAC	TGG	ATT	GTG
	Arg	His	Ser	Leu	Tyr	Val.	Asp	Phe	Ser	Asp	Val	Gly	Trp	Asn	Asp	Trp	Ile	Val
50			392					L407		~~			1422	m-c	~~	on the		1437
	Ala	Pro	Pro	Gly	TVY	Gln	Ala	Phe	TVI	CVS	His	Glv	Aso	CVS	Pro	Phe	CCA Pro	Leu
					145	2				146	7				148			
	GC	T GA	CCA	C CI	C AA	CTC	A AC	C AA	C CA	ፓ GC s እነ	C AI	T GI e Va	G CA	GAC n Tr	X CI	.G G7 eu Va	.C AA al A≤	T TCI n Ser
<i>55</i>	AL	a AS	א עז	عدد	באני	41 26		-				u						

	1497 1512 1527 1542
	GIC AAT TOO AGT ATO OOC AAA GOO TGT TGT GIG OOC ACT GAA CIG AGT GOO ATO Val Asn Ser Ser Ile Pro Lys Ala Cys Cys Val Pro Thr Glu Leu Ser Ala Ile
5	
	1557 1572 1587 TOC ATG CTG TAC CTG GAT GAG TAT GAT AAG GTG GTA CTG AAA AAT TAT CAG GAG
	Ser MET Leu Tyr Leu Asp Glu Tyr Asp Lys Val Val Leu Lys Asn Tyr Gln Glu
10	1602 1617 (408) 1636 1646 1656
	ATG GTA GTA GAG GGA TGT GGG TGC CGC TGAGATCAGG CAGTCCTTGA GGATAGACAG MET Val Val Glu Gly Cys Gly Cys Arg
	1666 1676 1686 1696 1706 1716 1726
15	ATATACACAC CACACACACA CACCACATAC ACCACACACA
	1736 1746 1756 1766 1776 1786 1796
	ACAGACTECT TOCTTATAGC TEGACTITTA TITAAAAAAA AAAAAAAAA AATEGAAAAA ATCOCTAAAC
20	
	1806 1816 1826 1836 1846 1856 1866 ATTCACCITG ACCITATITA TGACITIACS TGCAAATGIT TTGACCATAT TGATCATATA TTTTGACAAA
	MICACCITO ACCITATINA IGACITIMO IGARAGOII 11-20-22-2-2-2-2-2-2-2-2-2-2-2-2-2-2-2-2
25	1876 1886 1896 1906 1916 1926 1936
	ATATATTAT AACIACTAT TAAAAGAAA AAATAAAATG AGTCATTATT TTAAAAAAA AAAAAAAACT
30	
30	1946 CIAGAGIOGA OGGAATIC,
	wherein said process comprises the following steps:
35	a) screening of a gene library constructed from U-2 OS derived DNA or cDNA with a labelled bBMP-2 fragment
•	by hybridization,
	b) isolating positive clones, andc) isolating the DNA-inserts from said clones.
40	12. The process according to claim 11, wherein the gene encodes human BMP-4 having the amino acid sequence
••	given in claim 11.
	13. A process for the preparation of a gene encoding a protein exhibiting properties of BMP-4 and comprising a DNA
45	sequence:
	a) which differs from a DNA sequence of claim 11 in codon sequence due to the degeneracy of the genetic code;
	b) which hybridizes with DNA sequence of claim 11 or section a), above; or

50

wherein said process comprises standard techniques of molecular biology.

in changes in the peptide sequence or not,

- 14. The process according to claim 13, wherein the DNA sequence is a genomic DNA sequence.
- 15. The process according to claim 13, wherein the DNA sequence is a cDNA sequence.
 - **16.** A vector containing the gene or DNA sequence prepared according to any one of claims 1 to 15 in operative association with an expression control sequence.

c) represents a fragment, allelic or other variation of a DNA sequence of claim 11, whether said variation results

- 17. A cell transformed with a vector of claim 16.
- 18. The cell of claim 17 which is a mammalian cell, a bacterial cell, an insect cell or a yeast cell.
- 5 19. The cell of claim 18 which is a CHO cell.

10

15

25

35

40

45

50

55

- 20. A process for the preparation of a protein exhibiting properties of BMP-2, wherein said process comprises the steps of culturing in a suitable culture medium a cell transformed with an expression vector comprising a gene or a DNA sequence prepared according to any one of claims 1 to 10, and recovering said protein from said culture medium.
- 21. A process for the preparation of a protein exhibiting properties of BMP-4, wherein said process comprises the steps of culturing in a suitable culture medium a cell transformed with an expression vector comprising a gene or a DNA sequence prepared according to any one of claims 11 to 15, and recovering said protein from said culture medium.
- 22. A process for producing a protein exhibiting properties of BMP-2 or BMP-4, comprising the steps of culturing in a suitable culture medium the cell of claim 17 and isolating said protein from said culture medium.
- 20 23. A process for the preparation of a pharmaceutical composition comprising combining the proteins prepared according to any one of claims 20 to 22, individually or in combination with a pharmaceutically acceptable vehicle.
 - 24. The process according to claim 23, wherein said pharmaceutical composition further comprises a matrix capable of delivering the composition to the site of the bone or cartilage defect and providing a structure for inducing bone or cartilage formation.
 - 25. The process according to claim 24, wherein said matrix comprises hydroxyapatite, collagen, polylactic acid or tricalcium phosphate.
- 30 26. Use of a protein prepared according to any one of claims 20 to 22, individually or in combination, for the preparation of a pharmaceutical composition for inducing bone or cartilage formation.

Patentansprüche

Patentansprüche für folgende Vertragsstaaten : BE, CH, DE, FR, GB, IT, LI, LU, NL, SE

1. Menschliches BMP-2 codierendes Gen, umfassend die nachfolgende DNA-Sequenz:

	GTCGACT	10		_	0	1001		30	ome e		40		C 3 3	5(
	GICGACI	ICIA	GAGTG	1616	TC	AGCA	CTT	66 (CTGC	GGA	CII	CII	GAA	CIIC	3	
5	CAGGGAG	60 TAKE	-		0	0007		80 TC -		ጎርርጥ	90. GCC		ccc	100	_	
	CAGGGAG	JANI	AACII	GCGC	.д С								GCC			
	CGGAGC	110 TGC	ጥጥርርር			CGAG			cccc		140 TCC		ירכיתי	150 CGG	-	
10	COOROCC		11000													
	ommo o o o	160														
	CTTGCC	CGAC	ACTGA	GACG	C T	GTTC	CCA	GC (G'I'G <i>I</i>	AAAA	GAG	AGA	CTG	CGCC	j	
		210						30			240			250	_	
15	GCCGGC	ACCC	GGGAG	AAGG	A G	GAGO	CAA	AG .	LAAA	AGGA	ACG	GAC	TTA	CGG!	C	
		260		27							290			300		
	CCTTGC	GCCA	GGTCC	TTTG	A C	CAG	AGTT	TT	TCC	ATGT	GGA	CGC	TCT	TTC	A	
20		310		32	0		3	30			340			350)	
	ATGGAC	STGT	CCCC	CGTG	C T	TCT	raga	CG	GAC	rgcg	GTC	TCC	TAA	AGG:	r	
_	··							_						-		
-	(1 CGACC: ATI		- : GCC 0		0 ~ ~	·	 ai m	- -	38 Th. C	35 ~ m	~ ~	~ ~	m ~	~ 40 ~ ~	00	T
25	CACC AT	e end	GCC G	GG AC	$x \alpha$		FÎ CI 75 Le	T C. N La	EA GO	E T	e ci	rg Ci	m a	x a	S G	IC
25	CACC AT	G GTG T Val	Ala G	GG AC	$x \alpha$	sc N	/s Le	T C. N L	EA GO	E T	en Te GC	eu Le	T C	x a	S G	rc al
25	CTC CTG	G GTG T Val 4 GGC G	Ala G	GG AC	C CC	cic ig Ci	S Le 430 GIT	n L	TA GC PU A PAG	os Ti la La	eec m Fe	445	eu Pi AGG	C CI TO GI	KG G. Ln Va	ecc al
25	CACC AT	G GTG T Val 4 GGC G	Ala G	GG AC	C CC	cic ig Ci	S Le 430 GIT	n L	TA GC PU A PAG	os Ti la La	eec m Fe	445	eu Pi AGG	C CI TO GI	KG G. Ln Va	ecc al
	CTC CTG (Leu Leu 460	G GTG T Val GGC G Gly G	Ala o 315 SSC GOO 31y Ala	GG AC	GGC Gly	ren cic	7S Le 430 GIT Val	au La CCG Pro	EA GAG Glu 490	cre Leu	CGC CGC	445 CGC Arg	eu Pi AGG Arg	AAG Lys 505	AG G In Va TIC Phe	ed GCG Ala
	CTC CTG (Leu Leu CCC CCC CCC CCC CCC CCC CCC CCC CCC C	G GTG T Val GGC G Gly G	Ala o	GG ACT Ala 475	eec ety	LCY Terr CLC	430 GIT Val	cas Le	GAG GLU 490	CIG Leu	GAC GLY	445 CGC Arg	AGG Arg	AAG Lys 505	YG G Ln Va TTC Phe	ecc Ala
30	CTC CTG (Leu Leu 460	G GTG T Val GGC G Gly G	Ala o	GG ACT Ala 475	eec ety	LCY Terr CLC	430 GIT Val	cas Le	GAG GLU 490	CIG Leu	GAC GLY	445 CGC Arg	AGG Arg	AAG Lys 505	YG G Ln Va TTC Phe	ecc Ala
	CTC CTG (Leu Leu (460) GCG GCG (Ala Ala (G GTG T Val GGC G Gly G TGG T Ser S	Ala o	GG AC Ily Th GCT Ala 475 CGC	eec ely	CIC Leu TCA Ser 535	430 GIT Val TCC Ser	cas Le Cos Pro Cas Gln	GAG GLU 490 CC Pro	CIG Leu TCI Ser	GGC Gly GAC Asp 550	445 CGC Arg GAG Glu	AGG Arg GIC Val	AAG Lys 505 CTG Leu	AG G. TTC The AGC Ser	GAG GAU 565
30	CTC CTG (Leu Leu 460 CCC CCC Ala Ala Ala	G GTG T Val GGC G Gly G TGG T Ser S 520 TTG G	Ala o	GG AC Hy Th GCT Ala 475 CCC Arg	GGC Gly CCC Pro	CTC Leu TCA Ser 535 ATG	430 GIT Val TCC Ser	cos Pro CAS Gln	GAG GLU 490 CCC PTO	CIG Leu TCT Ser	GGC Gly GAC Asp 550 CAG	445 CGC Arg GAG Glu	AGG Arg GIC Val	AAG Lys 505 CIG Leu	TIC Phe AGC	GCG Ala GAG Glu 565 AGC
30	CTC CTG (Leu Leu (460) GCG GCG (Ala Ala (G GTG T Val GGC G Gly G TGG T Ser S 520 TTG G	Ala o	GG AC Hy Th GCT Ala 475 CCC Arg	GGC Gly CCC Pro	CTC Leu TCA Ser 535 ATG	430 GIT Val TCC Ser	cos Pro CAS Gln	GAG GLU 490 CCC PTO	CIG Leu TCT Ser	GGC Gly GAC Asp 550 CAG	445 CGC Arg GAG Glu	AGG Arg GIC Val	AAG Lys 505 CIG Leu	TIC Phe AGC	GCG Ala GAG Glu 565 AGC
30	CTC CTG (Leu Leu 460 GCG GCG Ala Ala Ala TTC GAG Phe Glu	G GTG T Val GGC G Gly G TGG T Ser S 520 TTG C Leu A	Ala 6 SEC SCC STY Ala SEC SCC SEC STY ATT TOS SCC SEC STY ATT TOS SCC TOS TOS SCC TOS TOS TOS TOS TOS TOS TOS T	GG AC Hy Tr GCT Ala 475 CCC Arg	GGC Gly CCC Pro AGC Ser	CIC Leu TCA Ser 535 AIG MET	430 GIT Val TCC Ser TTC Phe	COG Pro CAG Gln GGC Gly 595	GAG GLU 490 CCC PTO CIG	CIG Leu TCT Ser AAA Lys	GGC Gly GAC Asp 550 CAG Gln	445 CGC Arg GAG Glu AGA Arg	AGG Arg GIC Val CCC Pro 610	AAG Lys 505 CTG Leu ACC	TIC Phe AGC Ser	GAG GAU 565 AGC Ser
30	CTC CTG (Leu Leu CCC GCC GCC GCC GCC GCC GCC GCC GCC GC	G GTG T Val GGC G Gly G TGG T Ser S 520 TTG C Leu A	Ala 6 SEC GOO SET GLY ATG LET SEC GTO SET GLY SEC GTO SEC G	GG ACI GCT Ala 475 CCC Arg	esc ser	CIC Leu TCA Ser SIS AIG MET	430 GIT Val TCC Ser TTC Phe	COG Pro CAG Gln GGC Gly 595 CTA	GAG GLU 490 CCC PTO CIG Leu GAC	CIG Leu TCT Ser AAA Lys	GGC Gly GAC Asp S50 CAG Gln	445 CGC Arg GAG Glu AGA Arg	AGG Arg GIC Val CCC Pro 610 AGG	AAG Lys 505 CTG Leu ACC Thr	TTC Phe AGC Ser CCC Pro	GAG GAU 565 AGC Ser
30	CTC CTG (Leu Leu 460 GCG GCG Ala Ala Ala TTC GAG Phe Glu	G GTG T Val GGC G Gly G TGG T Ser S 520 TTG C Leu A	Ala 6 SEC GOO SET GLY ATG LET SEC GTO SET GLY SEC GTO SEC G	GG ACI GCT Ala 475 CCC Arg	esc ser	CIC Leu TCA Ser SIS AIG MET	430 GIT Val TCC Ser TTC Phe	COG Pro CAG Gln GGC Gly 595 CTA	GAG GLU 490 CCC PTO CIG Leu GAC	CIG Leu TCT Ser AAA Lys	GGC Gly GAC Asp S50 CAG Gln	445 CGC Arg GAG Glu AGA Arg	AGG Arg GIC Val CCC Pro 610 AGG	AAG Lys 505 CTG Leu ACC Thr	TTC Phe AGC Ser CCC Pro	GAG GAU 565 AGC Ser

5	625 CAG CCG Gln Pro	GGC								TTG							
10	AAC ACT Asn Thr																
	730 AGT GGG Ser Gly	AAA Lys	ACA Thr	ACC Thr	745 CGG Arg	AGA Arg	TTC Phe	TTC Phe	TTT Phe	760 AAT Asn	TIA Leu	AGT Ser	TCT Ser	ATC Ile	775 CC Pro	ACG Thr	GAG Glu
15	GAG TIT	790 ATC	ACC Thr	TCA Ser	GCA Ala	GAG Glu	805 CIT Leu	Gln CAG	GTT Val	TTC Phe	CGA Arg	820 GAA Glu	CAG Gln	ATG MET	CAA Gln	GAT Asp	835 GCT Ala
20	TTA GGI	A AAC 7 Asn	AAT Asn	850 AGC Ser	AGT Ser	TIC Phe	CAT His	CAC His	865 CGA Arg	ATT	AAT Asn	ATT Ile	TAT Tyr	880 GAA Glu	ATC Ile	ATA Ile	aaa Lys
25	895 CCT GCI Pro Ala	A ACA	GCC Ala	AAC Asn	TOG Ser	910 AAA Lys	TTC	e Pro	GTG Val	ACC Thr	925 AGT Ser	CIT Leu	TIG Leu	GAC Asp	ACC Thr	940 AGG Arg	TTG Leu
30	GTG AA	r CAG n Glr	955 TAA TSA	GCA	AGC Ser	ACG Arg	TCG	970 GAA Glu	AGT	TTT Phe	GAT Asp	GTC Val	985 ACC Thr	ccc Pro	GCT Ala	GTG Val	ATG MET
35	1000 CCG TG Arg Tr	p Thr	Ala	CAG	Glv	CAC His	Ala	AAC ASD	CAT	Gly	TIC	Val	Val	GAA Glu	1045 GTG Val	GCC	CAC His
	TTG GA	1060 G GAC U Glu	AAA	CAA	GGI	GIC	1075	Aag	AGA	CAI	gm	1090 'AGG	ATA	AGC) AGG	TCT	1105 TIG Leu
40	CAC CA His Gl	A GAI n Asj	GAA Glu	1120 CAC His	: AGC	Trp	TC: Ser	CAC	1135 ATA Ile	AGG	CP Pro	TIC Leu	CIA	1150 GIA Val	ACI	TII Phe	GC Gly
45	lle CAT GA His As	T GG	A AAA / Lys	GC Gly	CAI His	1180 C CCI S Pro	CIX	C CAC 1 His	: AAA : Lys	AG? Arg	1195 A GAA G Glu	AAA	ogi Arg	CAA Glr	GCC Ala	1210 : AAA : Lys	CAC
50	AAA CA Lys Gl	n Arg	1225 3 AA/ 3 Lys	α	CTI Lev	AAC 1 Lys	TO Se	1240 C AGO r Ser	TGI	AAC Lys	AGA	A CAC	1255 CCI Pro	TTC	TAC Tyr	GIG Val	GAC Asp
55	1270 TTC A Phe S	GT GA er As	c GI p Va	ı el	128 G TG y Tr	G AA	T GA n As	C TG TT q	G AT p Il	130 T GI e Va	G GC	T CC a Pr	c co o Pr	G GG o Gl	131 G TA Y Ty:	r CA	c ccc s Ala

5		TAC		CAC			IGC Cys		TTT			GCT		CAT			TCC	
10				ATT			ACG Thr		ಡ್					TCI				
10	GCA					ACA	1450 GAA Glu	crc			ATC					CIT		GAG
15			AAG		GIA		AAG Lys	AAC					GIT					cc cc cc
20	TCT	Arg CCC (39	TAG.		553 SCA 2		15(TAAA'		ATAA	157: ATAT			1583 CATA	TAT	_	593 ING .	AAAA	1603 AAGAAA
25	AAA	A													•			

- 3. Gen, das ein Protein codiert, das Eigenschaften von menschlichem BMP-2 zeigt, und eine DNA-Sequenz umfaßt, 30 die:
 - (a) sich in der Codonsequenz infolge der Degeneriertheit des genetischen Codes von einer DNA-Sequenz nach Anspruch 1 unterscheidet;
 - (b) mit einer DNA-Sequenz nach Anspruch 1 oder nach vorstehendem Absatz (a) hybridisiert; oder
 - (c) ein Fragment, eine allelische oder eine andere Variation einer DNA-Sequenz nach Anspruch 1 darstellt, unabhängig davon, ob die Variation zu Änderungen in der Peptidsequenz führt oder nicht.
 - 4. DNA-Sequenz nach Anspruch 3, dadurch gekennzeichnet, daß sie eine genomische DNA-Sequenz ist.
- 40 5. DNA-Sequenz nach Anspruch 3, dadurch gekennzeichnet, daß sie eine cDNA-Sequenz ist.
 - 6. Rinder-BMP-2 codierendes Gen, umfassend die nachfolgende DNA-Sequenz:

50

45

35

5	(1) GGC C G H	lac . I	GAT D	GGG G	15 AAA K	GGA G	CAC	CCT P	CTC	30 CAC H	AGA R	AGA R	GAA E	AAG K	45 CGG R
	CAA (GCA .	AAA K	CAC H	60 AAA K	CAG Q	CGG R	AAA K	CGC R	75 CTC L	AAG K	TCC S	AGC S	TGT C	90 AAG K
10	AGA (CCT P	TTA L	105 TAT Y	GTG V	GAC D	TTC F	AGT S	120 GAT D	g t g V	GGG G	TGG W	AAT N	135 GAC D
15	TGG 1		GTT V	GCA A	150 CCG P	CCG P	GGG G	TAT Y	CAT H	165 GCC A	TIT	TAC Y	TGC C	CAT H	180 GGG G
20	GAG 1	TGC C	CCT P	TTT F	P	CTG L	GCC A	GAT D	CAC	210 CTT L	AAC N	TCC S	ACG T	AAT N	225 CAT H
as.	GCC :	ATT I	CTC V	CAA Q	240 ACT	CTĠ	GTC V	AAC N	TCA S	255 GTT V	AAC N	S	K	I,	₽
25	AAG K	GCA	TGC	TGT C	385 GTC V	CCA P	ACA	GAG	CTC	300 AGC S	GCC		TCC S		
30		CTT L	GAT D		330 AAT N	GAG	AAG K	GTG V	GTA V	345 TTA L	AAG	AAC N	TAT Y	CAG O	360 GAC D
35	ATG M	GTT V	GTC V	GAG E	GGT	TGT	GGG G	TGI	CGI	9) 'TAG			AAAT		07 .TA
40	TAAAT		17 FA T.	TATA		27 CA T	TAGA	43 AAAA		CAAA	441 AAAI		AAGT	457 TGAC	
45	YCIII		67 AT T	TCCC	47 AATO		GACT	48 TTAT		.TGG?	491 ATG		TGGA	507 GAAA	
	AAGAA		17 CA C	AGCI		27 ET G.	AAAA		7 A TI	TATA	54° TCT?		GAAA	557 AGAA	
50	GTTGG		67 AA C	AAAT	_	77 PT A	ATCA	58 GAGA		TATT	,				

- 7. Gen, das Rinder-BMP-2 codiert, das die Aminosäuresequenz von Anspruch 6 enthält.
- 8. Gen, das ein Protein codiert, das Eigenschaften von Rinder-BMP-2 zeigt, und DNA-Sequenzen umfaßt, die:
 - (a) sich in der Codonsequenz infolge der Degeneriertheit des genetischen Codes von einer DNA-Sequenz nach Anspruch 7 unterscheiden;

(b) mit einer DNA-Sequenz nach Anspruch 7 oder nach vorstehendem Absatz (a) hybridisieren; oder
(c) Fragmente, allelische oder andere Variationen einer DNA-Sequenz nach Anspruch 7 darstellen, unabhängig davon, ob die Variationen zu Änderungen in der Peptidsequenz führen oder nicht.
9. DNA-Sequenz nach Anspruch 8, dadurch gekennzeichnet, daß sie eine genomische DNA-Sequenz ist.
10. DNA-Sequenz nach Anspruch 8, dadurch gekennzeichnet, daß sie eine cDNA-Sequenz ist.
11. Menschliches BMP-4 codierendes Gen, umfassend die nachfolgende DNA-Sequenz:
10. 20. 30. 40. 50.
CTCTAGAGGG CAGAGGAGGA GGGAGGGAGG GAAGGAGCCCGGC

		CCG	GAA	60 GCT <i>I</i>		TGA	7 G TG I	o G G			80 .GC	TGAC	GGA	90 CGC	GAG	CCT	100 GAG	
5		CGC	ccc	110 TGC		TCC	12 GGCT		GTA:		.30 .GC	TTGT		140 CCC		GGG.	150 ATT	
10		cco	STCC	160 AAG		TCT	17 CGAG		TGC			ACAC		190 CCG		CTC	200 GCC0	_
		AGG	STTC	21(ACT(ACC	22 GTTC	0 2A G			30 CA	GGAG		240 CTG		GCG	25 AGC	
15		CGC	TAC	260 TGC		GAC	27 CTAI	o NG G	AGC		80 CC	GTAC		290 CAT		GAG	300 CAA	
		GCA	CTG	310 CTG		CTT	32 CCCI	O CGA			30 CA	GCA		340 TGT		AGA	350 PTG0	
20		сто	STCA	360 AGA		ATG	37 GACI	O CG T		3 PATA				390 TTC		CAA	400 GAC	
25			ATG	AT:	r cc	T												
30	GGT Gly	417 AAC Asn	CGA.	ATG MET	cic Leu	ATG MET	432 GIC Val	GTT	TTA	TTA	TGC	CAA	GIC	CTG	CTA Leu	GGA	462 GGC Gly	GCG Ala
:	AGC	CAT	GCT	477 AGT	TIG	ATA	OCT Pro	GAG	492 ACG	GGG	AAG	AAA	AAA	507 GTC	GCC	GAG	ATT	CAG
35	522 GGC Gly	CAC His	GOG Ala	GGA Gly	CJA CCY	537 ŒC Arg	CGC Arg	TCA Ser	GGG GGG	CAG Gln	552 AGC Ser	CAT His	GAG Glu	CIC Leu	CIG Leu	567 CCG Arg	GAC Asp	TTC Phe
40	GAG Glu	GCG Ala	582 ACA Thr	CIT Leu	CIG Leu	CAG Gln	ATG MET	597 TIT Phe	GGG Gly	CIG Leu	OGC Arg	OGC Arg	612 CGC Arg	ccc Pro	CAG CAG	CCT Pro	AGC Ser	627 AAG Lys
45	AGT Ser	GCC Ala	GIC Val	ATT Ile	642	GAC Asp	TAC Tyr	ATG MET	CCG Arg	657 GAT Asp	CIT Leu	TAC Tyr	CGG Arg	CTT Leu	672 CAG Gln	TCT Ser	GGG Gly	GAG Glu
50	GAG Glu	687 GAG Glu	GAA Glu	GAG Glu	CAG Gln	ATC Ile	702 CAC His	AGC Ser	ACT Thr	GGT Gly	CTT Leu	717 GAG Glu	TAT Tyr	CCT Pro	GAG Glu	CGC Ar g	732 CCG Pro	GCC Ala
50	AGC Set	COC Arg	GCC Ala	747 AAC ASD	ACC	GIG Val	AGG Arg	AGC Ser	762 TIC Phe	CAC	CAC His	GAA Glu	GAA Glu	777 CAT His	CTG	GAG -Glu	AAC Asn	ATC Ile

	792					807					822					837		
	∞ A	GGG	∞	agt	G4A	AAC	TCT	GCT	TTT	α	TTC	CIC	TIT	AAC	CIC	AGC	AGC	ATC
5	Pro	Gly	Thr	Ser	Glu	يجلا	Ser	Ala	Pt.e	Arg	Phe	Leu	Phe	Asn	Leu	Ser	Ser	Ile
			050															205
	Д	C3C	852	C3C	~	3000	~~~	867	~~	C3C	~	~~	882		~~	~~	س د	897
														Phe				GTG Val
10					•				ALG	676		AL 9						•
					912					927					942			
														AAC				
	Asp	Gln	Gly	Pro	Asp	dz_L	Glu	yrg	Gly	Phe	His	yrg	Ile	Asn	Ile	Tyr	Glu	Val
15		957																
	ATYG		\sim	رتب	CC3	633	972	~~~	~		~~	987	3000	3.03	~3	-	1002	GAC
																		Asp
		-,-					•	••••	110	CI		264			9			עביי
20				1017					1032					L047				
20														TIT				
	m	yià	Leu	Val	His	His	yzu	Val	Ţħ.	Arg	طتي	Glu	Thr	Phe	yżb	Val	Ser	bio
	1062					1077					1092						•	
			CTT	csc			œ	G)G	236			220	ጥልጥ	ccc		1107	2000	GAG
25	Ala	Val	Leu	Arq	Tro	Thr	Ara	Glu	Lvs	Gln	Pm	Asn	Tyr	Glv	Leu	Ala	Tle	Glu
				•	-•		5						-1-	1				
•			1122					1137					152					L167
	GIG	ACI	CAC	CIC	CAT	CAG	ACT	œ	\overline{w}	CYC	CYC	GGC	CAG	CAT	GIC	AGG	ATT.	AGC
30	val	шш	nis	Teu	HIS	GID	.mr	Arg	Inr	HIS	GIN	GIY	GTJ	His	Val	Arg	He	Ser
					182					197					212			
														ಜ್				
	My	Ser	TEU	PIO	GIN	σλ	Ser	GIĀ	ASN	up	ATA	GTV	Leu	Arg	Pro	Leu	Leu	Val
35	. 1	227				1	242					L257				3	272	
	ACC	TTT	GGC	CAT	CAT			GGC	CAT	GCC			CGA	ಹರ	ಯ			AAG
	Thr	Phe	Gly	His	Asp	Gly	Arg	Gly	His	Ala	Leu	Thr	Arg	Arg	Arg	Arg	Ala	Lys
			,	207														
40	€	ACC.		287	CD/TP	C3 C	TY**A		1302	~~	300	220		1317	220	220	m-c	ထင
	Arq	Ser	Pro	Lvs	His	His	Ser	Gln	Arti	Ala	Arer	Tys	IVS	YZU YVI	TVS	ASD	CVS	Arg
				-1-				<u> </u>	9		9	-,-			<i>ح</i> ر پ		٥٢٥	
	1332					1347					1362					L377		
45	œc	CAC	TOG	cic	TAT	CIG	GAC	TIC	AGC	GAT	GIG.	GGC	TGG	AAT	GAC	TGG	ATT	GIG
45	Arg	HIS	Ser	Leu	Tyr	Val	Asp	Phe	Ser	yzb	Val	Gly	dır	Asn	Asp	ITP	Ile	Val
			1.392				,	1407					1422					1437
	ccc			GGC	TAC	CAG			TAC	TGC	ĊAT			MGC.	9	بلبليل		CIG
														CVS				
50																		
	cr	יוף פו	ער כי	ر س	14! ^ 2:		יג בי	~ ≥,	יר כי	146 T G		ىتى بات	ים בי	G 20	148 77	-	تد ب	T TCT
																		n Ser

5	1497 1512 1527 1542 GTC AAT TOO AGT ATC OOM AAA GOO TGT TGT GTG OOM ACT GAA CTG AGT GOO ATC Val Asn Ser Ser Ile Pro Lys Ala Cys Cys Val Pro Thr Glu Leu Ser Ala Ile
10	1557 1572 1587 TCC ATG CTG TAC CTG GAT GAG TAT GAT AAG GTG GTA CTG AAA AAT TAT CAG GAG Ser MET Leu Tyr Leu Asp Glu Tyr Asp Lys Val Val Leu Lys Asn Tyr Gln Glu 1602 1617 (408) 1636 1646 1656
15	ATG GTA GTA GAG GGA TGT GGG TGC CGC TGAGATCAGG CAGTCCTTGA GGATAGACAG MET Val Val Glu Gly Cys Gly Cys Arg
15	
	1666 1676 1686 1696 1706 ATATACACAC CACACACA CACCACATAC ACCACACACA
20	1716 1726 1736 1746 1756 TCCACTCACC CACACACTAC ACAGACTGCT TCCTTATAGC TGGACTTTTA
25	1766 1776 1786 1796 1806 TTTAAAAAA AAAAAAAAA AATGGAAAAA ATCCCTAAAC ATTCACCTTG 1816 1826 1836 1846 1856
	ACCTTATTTA TGACTTTACG TGCAAATGTT TTGACCATAT TGATCATATA 1866 1876 1886 1896 1906
30	TTTTGACAAA ATATATTTAT AACTACGTAT TAAAAGAAAA AAATAAAATG 1916 1926 1936 1946
35	AGTCATTATT TTAAAAAAAA AAAAAAAACT CTAGAGTCGA CGGAATTC
	 12. Gen, das menschliches BMP-4 codiert, das die in Anspruch 11 angegebene Aminosäuresequenz aufweist. 13. Gen, das ein Protein codiert, das Eigenschaften von BMP-4 zeigt, und eine DNA-Sequenz umfaßt, die:
40	(a) sich in der Codonsequenz infolge der Degeneriertheit des genetischen Codes von einer DNA-sequenz nach Anspruch 11 unterscheidet;
	(b) mit einer DNA-Sequenz nach Anspruch 11 oder nach vorstehendem Absatz (a) hybridisiert; oder
45	(c) ein Fragment, eine allelische oder eine andere Variation einer DNA-Sequenz nach Anspruch 11 darstellt, unabhängig davon, ob die Variation zu Änderungen in der Peptidsequenz führt oder nicht.
	14. DNA-Sequenz nach Anspruch 13, dadurch gekennzeichnet, daß sie eine genomische DNA-Sequenz ist.
50	15. DNA-Sequenz nach Anspruch 13, dadurch gekennzeichnet, daß sie eine cDNA-Sequenz ist.
	16. Vektor, enthaltend das Gen oder die DNA-Sequenz nach einem der Ansprüche 1 bis 15 in einer funktionellen Verbindung mit einer Expressions-Kontrollsequenz.
55	17. Zelle, dadurch gekennzeichnet, daß sie mit einem Vektor nach Anspruch 16 transformiert ist.

18. Zelle nach Anspruch 17, dadurch gekennzeichnet, daß sie eine Säugerzelle, eine Bakterienzelle, eine Insekten-

zelle oder eine Hefezelle ist.

- 19. Zelle nach Anspruch 18, dadurch gekennzeichnet, daß sie eine CHO-Zelle ist.
- 20. Protein, das Eigenschaften von BMP-2 aufweist, das durch ein Gen oder eine DNA-Sequenz nach einem der Ansprüche 1 bis 10 codiert ist. -

5

- 21. Protein, das Eigenschaften von BMP-2 aufweist, das erhältlich ist durch die Schritte
 - Züchten einer mit einem Expressionsvektor transformierten Zelle in einem geeigneten Kulturmedium, wobei der Vektor ein Gen oder eine DNA-Sequenz nach einem der Ansprüche 1 bis 10 umfaßt, und

10

Gewinnen des Proteins aus dem Kulturmedium.

22. Protein, das Eigenschaften von BMP-4 aufweist, das durch ein Gen oder eine DNA-Sequenz nach einem der Ansprüche 11 bis 15 codiert ist.

15

- 23. Protein, das Eigenschaften von BMP-4 aufweist, das erhältlich ist durch die Schritte
 - Züchten einer mit einem Expresionsvektor transformierten Zelle in einem geeigneten Kulturmedium, wobei der Vektor ein Gen oder eine DNA-Sequenz nach einem der Ansprüche 11 bis 15 umfaßt und

20

- Isolieren des Proteins aus dem Kulturmedium.
- 24. Verfahren zur Herstellung des Proteins nach Anspruch 21 oder 23, umfassend die Schritte

25

- Züchten der Zelle nach Anspruch 17 in einem geeigneten Kulturmedium und
- Gewinnen des Proteins aus dem Kulturmedium.

30

- 25. Arzneimittel, dadurch gekennzeichnet, daß es, einzeln oder in Kombination, die Proteine nach einem der Ansprüche 20 bis 23 und einen pharmakologisch verträglichen Träger umfaßt.
- 26. Arzneimittel nach Anspruch 25, dadurch gekennzeichnet, daß es ferner eine Matrix umfaßt, die fähig ist, das Arzneimittel an die Stelle des Knochen- oder Knorpelschadens zu liefern und eine Struktur zur Induktion der Knochen- oder Knorpelbildung bereitzustellen.

35

- 27. Arzneimittel nach Anspruch 26, dadurch gekennzeichnet, daß die Matrix Hydroxyapatit, Kollagen, Polyessigsäure oder Tricalciumphosphat umfaßt.
- 28. Verwendung des Proteins nach einem der Ansprüche 20 bis 23, einzeln oder in Kombination, zur Herstellung eines 40 Arzneimittels zur Induktion der Knochen- oder Knorpelbildung.

Patentansprüche für folgenden Vertragsstaat : AT

Verfahren zur Herstellung eines menschliches BMP-2 codierenden Gens, das die nachfolgende DNA-Sequenz umfaßt:

50

45

GTC	GACT								GG					GAA			
CAGO	GGAG	-	_	CTT			CCC			CGCC	GGT			GCC			
CGG!	AGC			CGC		_		_		CCGC				CCT		-	
CTT	GCC			TGAG			GTT			GTG				CTG			
GCC	GGC			GAG			GAG							:ATT			
CCT	rgc			TCC			CAG			TCC	\TGT			TCT		_	
ATG	GAC			CCG			TCT:			GACT	rgcg			TAA		-	
CGAC	CAI	ंद्र दा	G GC	∝ co la Gi	G AC	$x \alpha$	rg Cj SC II	sic /s L	M C	IA GO	E T	e C	න T අ ශ ය	er co	x c	S G	r C
CTC (CIG Leu	GGC Gly	415 GGC Gly	GCG Ala	GCT Ala	GGC GGC	CIC Leu	430 GTT Val	ccc Pro	GAG Glu	CIG Leu	GGC Gly	445 CGC Arg	AGG Arg	aag Lys	TIC Phe	GCG Ala
460 GGG (GOG Ala	TCG Ser	TCS Ser	GGC Gly	475 CGC Arg	œ Pro	TCA Ser	TCC Ser	CAG Gln	490	TCT Ser	GAC Asp	GAG Glu	GTC Val	505 CIG Leu	AGC Ser	GAG Glu
TIC	GAG	520 TTG	ŒG	CIG	crc	AGC	535 ATG	TTC	GGC	crg	AAA	550 CAG	AGA	α	ACC	ಱ	565 AGC
acc :	GYC	ထေ	GIG	580 GTG	œ	∞	TAC	ATG	595 CTA	GAC	CTG	TAT	œc	610 AGG	CAC	TOG	GGT
	CAGO CGGA CTTC GCCC CCTT ATGC CTC Leu 460 GCG Ala TTC Phe	CAGGGAGC CGGAGCC CTTGCC GCCGGCC ATGGAC ATGGA	CAGGGAGAAT CAGGGAGCCTGC 160 CTTGCCCGAC 210 GCCGGCACCC 260 CCTTGCGCCAC ATGGACGTGT ATGGACGTGT ATGGACGTGT ATGGACGTGT ATG GGC Leu Leu Gly 460 GCC GCC TCG Ala Ala Ser TCC GAG TTG Ehe Glu Leu AGG GAC GCC	GTCGACTCTA GA 60 CAGGGAGAAT AA 110 CGGAGCCTGC TT 160 CTTGCCCGAC AC 210 GCCGGCACCCC GG CCTTGCGCCA GG ATGGACGTGT CC ATGGACGTGT CC MET Val AI CTC CTG GGC GGC Leu Leu Gly Gly 460 GCG GCC TCG TCG Ala Ala Ser Ser TTC GAG TTG CGG Ehe Glu Leu Arg AGG GAC GCC GTG	CAGGGAGAAT AACTTO CAGGGAGCCTGC TTCGCC 160 CTTGCCCGAC ACTGAC 210 GCCGGCACCC GGGAGA 260 CCTTGCGCCA GGTCC ATGGACGTGT CCCCGC ATGGACGTGT CCCCGC ATG GTG GCC GCC MET Val Ala GI CTC CTG GGC GGC GCC Leu Leu Gly Gly Ala 460 GCC GCC TCC TCC GCC Ala Ala Ser Ser Gly TTC GAG TTG CGC CTG Ene Glu Leu Arg Leu AGG GAC GCC GTG GTG 580 AGG GAC GCC GTG GTG 580 AGG GAC GCC GTG GTG 580	GTCGACTCTA GAGTGTGTG 60 7 CAGGGAGAAT AACTTGCGC 110 12 CGGAGCCTGC TTCGCCATC 160 17 CTTGCCCGAC ACTGAGACG 210 22 GCCGGCACCC GGGAGAAGG 260 27 CCTTGCGCCA GGTCCTTTG 310 32 ATGGACGTGT CCCCGCGTG ATG GTG GCC GGG AC MET Val Ala Gly TI CTC CTG GGC GGC GCT Leu Leu Gly Gly Ala Ala 460 475 GCG GCG TCG TCG GGC CCC Ala Ala Ser Ser Gly Arg TTC GAG TTG CGG CTG CTC The Glu Leu Arg Leu Leu 580 AGG GAC GCC GTG GTG CCC	GTCGACTCTA GAGTGTGTGT C 60 70 CAGGGAGAAT AACTTGCGCA C 110 120 CGGAGCCTGC TTCGCCATCT C 160 170 CTTGCCCGAC ACTGAGACGC T 210 220 GCCGGCACCC GGGAGAAGGA G 260 270 CCTTGCGCCA GGTCCTTTGA C 310 320 ATGGACGTGT CCCCGCGTGC T CACC ATG GTG GCC GGG ACC CC MET Val Ala Gly Thr A CTC CTG GGC GGC GCG GCT GGC Leu Leu Gly Gly Ala Ala Gly 460 CCC GCG TCG TCG GGC CCC Ala Ala Ser Ser Gly Arg Pro TTC GAG TTG CGG CTG CTC AGC TTC GAG TTG CTG CTC AGC TTC GTG CTC AGC TTC GTG CTC CTC CTC AGC TTC GTG CTC CTC AGC TTC GTG CTC CTC AGC TTC GTG CTC CTC CTC AGC TTC GTG CTC CTC CTC AGC TTC CTC CTC CTC AGC TTC CTC CTC CTC CTC CTC AGC TTC CTC CTC CTC CTC CTC CTC CTC	CAGGGAGAAT AACTTGCGCA CCCCA 110 120 CGGAGCCTGC TTCGCCATCT CCGAC 160 170 CTTGCCCGAC ACTGAGACGC TGTTC 210 220 GCCGGCACCC GGGAGAAGGA GGAGC 260 270 CCTTGCGCCA GGTCCTTTGA CCAGA 310 320 ATGGACGTGT CCCCGCGTGC TTCTC (1) 370 CGACC ATG GTG GCC GGC ACC CCC TC MET Val Ala Gly Thr Arg Cy 415 CTC CTG GGC GGC GGC GGC CCC TC 160 475 CGC GGC TGC TGC GGC CTC 161 Leu Gly Gly Ala Ala Gly Leu 460 CGC GGC TGC TGC GGC CCC TCA Ala Ala Ser Ser Gly Arg Pro Ser 520 TTC GAG TTG GGG CTG CTC AGC ATG Phe Glu Leu Arg Leu Leu Ser MET AGG GAC GGC GTG GTG CCC CCC TAC AGG GAC GGC GTG GTG CCC CCC TAC	CAGGACTCTA GAGTGTGTGT CAGCACTT 60 70 CAGGGAGAAT AACTTGCGCA CCCCACTT 110 120 1 CGGAGCCTGC TTCGCCATCT CCGAGCCC 160 170 1 CTTGCCCGAC ACTGAGACGC TGTTCCCA 210 220 22 GCCGGCACCC GGGAGAAGGA GGAGGCAA 260 270 270 CCTTGCGCCA GGTCCTTTGA CCAGAGTT 310 320 320 ATGGACGTGT CCCCGCGTGC TTCTTAGA ATGGACGTGT CCCCGCGTGC TTCTTAGA (1) 370 CGACC ATG GTG GCC GGG ACC GGC TGT CC MET Val Ala Gly Thr Arg Cys Le 415 430 CTC CTG GGC GGC GGC GGC CTC GTT Leu Leu Gly Gly Ala Ala Gly Leu Val 460 475 CCC GCC GCC TCC TCC GCC CCC TCA TCC Ala Ala Ser Ser Gly Arg Pro Ser Ser TTC GAG TTG CGG CTG CTC AGC ATG TTC Phe Glu Leu Arg Leu Leu Ser MET Phe 580 AGG GAC GCC GTG GTG CCC CCC TAC ATG	GTCGACTCTA GAGTGTGTGT CAGCACTTGG 60 70 80 CAGGGAGAAT AACTTGCGCA CCCCACTTTG 110 120 130 CGGAGCCTGC TTCGCCATCT CCGAGCCCCA 160 170 180 CTTGCCCGAC ACTGAGACGC TGTTCCCAGC 210 220 230 GCCGGCACCC GGGAGAAGGA GGAGGCAAAG 260 270 280 CCTTGCGCCA GGTCCTTTGA CCAGAGTTTT ATGGACGTGT CCCCGCGTGC TTCTTAGACG (1) 370 CGACC ATG GTG GCC GGG ACC CGC TGT CTT CT CAGACC ATG GTG CGC GGC CTC GTT CCG Leu Leu Gly Gly Ala Ala Gly Thr Arg Cys Leu L 415 430 CTC CTG GCC GCC GCC GCC CCC TCA TCC CAGACC GCC GCC GCC CCC GTT CCCC GCC GCC CCC GTT CCCC GCC G	GTCGACTCTA GAGTGTGTGT CAGCACTTGG CTGG 60 70 80 CAGGGAGAAT AACTTGCGCA CCCCACTTTG CGCC 110 120 130 CGGAGCCTGC TTCGCCATCT CCGAGCCCCA CCGC 160 170 180 CTTGCCCGAC ACTGAGACGC TGTTCCCAGC GTGA 210 220 230 GCCGGCACCC GGGAGAAGGA GGAGGCAAAG AAAA CCTTGCGCCA GGTCCTTTGA CCAGAGTTTT TCCA 310 320 330 ATGGACGTGT CCCCGCGTGC TTCTTAGACG GACT (1) 370 38 CGACC ATG GTG GCC GGG ACC CGC TGT CTT CTA GC MET Val Ala Gly Thr Arg Cys Leu Leu Al 415 430 CTC CTG GGC GGC GCC GCC CTC GTT CCG GAG Leu Leu Gly Gly Ala Ala Gly Leu Val Pro Glu 460 475 490 GCG GCG TCG TCG GGC CCC CTC GTT CCG GAG Ala Ala Ser Ser Gly Arg Pro Ser Ser Gln Pro 520 535 TTC GAG TTG CGG CTG CTC AGC ATG TTC GGC CTG Phe Glu Leu Arg Leu Leu Ser MET Phe Gly Leu 580 595 ACG GAC GCC GTG GTG CCC CCC TAT CTA GAC	GTCGACTCTA GAGTGTGTGT CAGCACTTGG CTGGGGA 60 70 80 CAGGGAGAAT AACTTGCGCA CCCCACTTTG CGCCGGT 110 120 130 CGGAGCCTGC TTCGCCATCT CCGAGCCCCA CCGCCCC 160 170 180 CTTGCCCGAC ACTGAGACGC TGTTCCCAGC GTGAAAA 210 220 230 GCCGGCACCC GGGAGAAAGGA GGAGGCAAAG AAAAGGA 260 270 280 CCTTGCGCCA GGTCCTTTGA CCAGAGTTTT TCCATGT 310 320 330 ATGGACGTGT CCCCGCGGTGC TTCTTAGACG GACTGCG (1) 370 385 MET Val Ala Gly Thr Arg Cys Leu Leu Ala Leu 415 430 CTC CTG GGC GGC GGC GGC CTC GTT CTG GAG CTG Leu Leu Gly Gly Ala Ala Gly Leu Val Pro Glu Leu 460 475 490 GGG GGG TGG TGG GGC CTC GTT CTG GAG CTG Ala Ala Ser Ser Gly Arg Pro Ser Ser Gln Pro Ser TTC GAG TTG CGG CTG CTC AGC ATG TTC GGC CTG AAA Ahe Glu Leu Arg Leu Leu Ser MET Phe Gly Leu Lys 580 595 AGG GAC GGC GTG GTG CTC AGC CTG AGC CTG AGG GAC GGC GTG GTG CTC ATG CTA GAC CTG	GTCGACTCTA GAGTGTGTGT CAGCACTTGG CTGGGGACTT 60 70 80 90 CAGGGAGAAT AACTTGCGCA CCCCACTTTG CGCCGGTGCC 110 120 130 140 CGGAGCCTGC TTCGCCATCT CCGAGCCCCA CCGCCCCTCC 160 170 180 190 CTTGCCCGAC ACTGAGACGC TGTTCCCAGC GTGAAAAGAG 210 220 230 240 GCCGGCACCC GGGAGAAGGA GGAGGCAAAG AAAAGGAACG 260 270 280 290 CCTTGCGCCA GGTCCTTTGA CCAGAGTTTT TCCATGTGGA 310 320 330 340 ATGGACGTGT CCCCGCGTGC TTCTTAGACG GACTGCGGTC (1) 370 385 CGACC ATG GTG GCC GGG ACC GGC TGT CTT CTA GGC TTG CCAGACCT TGT VAI Ala Gly Thr Arg Cys Leu Leu Ala Leu Leu Leu Gly Gly Ala Ala Gly Leu Val Pro Glu Leu Gly 460 475 490 CCC CTG GGC TGG TGC CCC TCA TCC CAG CCC TCT GAC ALa Ala Ser Ser Gly Arg Pro Ser Ser Gln Pro Ser Asp 520 535 550 TTC GAG TTG CGG CTG CTC AGC ATG TTC GGC CTG AAA CAG TCC GTT CGG CAG CTG TTC CTG CGC CTG AAA CAG TCC GGC CTG ATG CTG CTG CTG CTG CAG CTG TTC CGG CTG AAA CAG TCC GGC CTG ATG CTG CGC CTG AAA CAG TCC GGC CTG AAA CAG TCC GGC CTG AAA CAG TCC CTG CGC CTG AAA CAG TCC CGG CTG CTG CTG CTG CTG CTG CTG CT	GTCGACTCTA GAGTGTGTGT CAGCACTTGG CTGGGGACTT CTT 60 70 80 90 CAGGGAGAAT AACTTGCGCA CCCCACTTTG CGCCGGTGCC TTT 110 120 130 140 CGGAGCCTGC TTCGCCATCT CCGAGCCCCA CCGCCCCTCC ACT 160 170 180 190 CTTGCCCGAC ACTGAGACGC TGTTCCCAGC GTGAAAAGAG AGA 210 220 230 240 GCCGGCACCC GGGAGAAGGA GGAGGCAAAG AAAAGGAACG GAC 260 270 280 290 CCTTGCGCCA GGTCCTTTGA CCAGAGTTTT TCCATGTGGA CGC 310 320 330 340 ATGGACGTGT CCCCGCGTGC TTCTTAGACG GACTGCGGTC TCC (1) 370 385 CGACC ATG GTG GCC GGG ACC CCC TGT CTT CTA GCC TTG CTG CCC MET Val Ala Gly Thr Arg Cys Leu Leu Ala Leu Leu Leu Leu Leu Gly Cly Ala Ala Gly Leu Val Pro Glu Leu Gly Arg 450 475 490 CGC GGC TCC TCC GGC CCC TCT CAC CAC CCC 1520 535 550 TTC CAC TTG CGC CTC CTC AGC ATG TTC CGC CTG AAA CAC AGA 250 595 CGC GAC CCC GTG GTG CCC CCC TTA TCC CAC CTG TAT CCC 150 595 CGC GAC CCC GTG GTG CCC CCC TTA TCC CAC CTG TAT CCC 150 595 CGC GAC CCC GTG GTG CCC CCC TTA TCC CAC CTG TAT CCC 150 595 CCC GAC CCC GTG GTG CCC CCC TTA TCC CAC CTG TAT CCC 150 595 CCC GAC CCC GTG GTG CCC CCC TTA TCC CAC CTG TAT CCC 150 595 CCC GAC CCC GTG GTG CCC CCC TTA CCC CTG TAT CCC 150 595 CCC GAC CCC GTG GTG CCC CCC TTA CCC CTG TAT CCC 150 595 CCC GAC CCC GTG GTG CCC CCC TTA CCC CTG TAT CCCC 150 595 CCC GAC CCC GTG GTG CCC CCC TTA CCC CTG TAT CCCC 150 595 CCC GAC CCC GTG GTG CCC CCC TTA CCC CTG TAT CCCC	CAGGGAGAAT AACTTGCGCA CAGGGAGAAT AACTTGCGCA CCCCACTTTG CGCCGGTGCC TTTGCC 110 120 130 140 CGGAGCCTGC TTCGCCATCT CCGAGCCCCA CCGCCCCTCC ACTCCT 160 170 180 190 CTTGCCCGAC ACTGAGACGC TGTTCCCAGC GTGAAAAGAG AGACTG 210 220 230 240 GCCGGCACCC GGGAGAAGGA GGAGGCAAAG AAAAGGAACG GACATT 260 270 280 290 CCTTGCCCGAC GGTCCTTTGA CCAGAGTTTT TCCATGTGGA CGCTCT 310 320 330 340 ATGGACGTGT CCCCGCGTGC TTCTTAGACG GACTGCGGTC TCCTAA ATGGACGTGT CCCCGCGTGC TTCTTAGACG GACTGCGGTC TCCTAA (1) 370 385 CGACC ATG GTG GCC GGG ACC CGC TGT CTT CTA GCC TTG CTG CTT CTC CTG GCC GGC ACG Leu Leu Gly Gly Ala Ala Gly Leu Val Pro Glu Leu Gly Arg Arg 450 475 490 CGC GCC TCC TCC GGC CCC CCC TCA TCC CAG CCC TCT CAC GAG GTC Ala Ala Ser Ser Gly Arg Pro Ser Ser Gln Pro Ser Asp Glu Val 520 535 550 TTC CAC TTG CGC CTC CTC AGC ATG TTC CGC CTG AAA CAC AGA CCC The Glu Leu Arg Leu Leu Ser MET Phe Gly Leu Lys Gln Arg Pro 580 595 610 AGG GAC GCC GTG GTG CTC AGC CTC TAT CCC AGC AGG CTC TAT CCC AGC TAC ATG CTG CTG TTT CTA ATG CTA CAC CTG TAT CCC AGG AGG CAC CTC TTT CTA CTC CAC CTT TAT CCC AGG TTC CAC GTG GTG CTC AGC ATG TTC CGC CTG AAA CAC AGA CCC TTC CAC GTG GTG CTC AGC ATG TTC CGC CTG AAA CAC AGA CCC TTC CAC GTG GTG CTC AGC ATG TTC CGC CTG AAA CAC AGA CCC TTC CAC GTG GTG CTC AGC ATG TTC CGC CTG AAA CAC AGA CCC TTC CAC GTG GTG CTC AGC ATG TTC CGC CTG AAA CAC AGA CCC TTC CAC GTG GTG CTC AGC ATG TTC CGC CTG AAA CAC AGA CCC TTC CAC GTG GTG CTC AGC ATG TTC CGC CTG AAA CAC AGA CCC TTC CAC GTG GTG CTC AGC ATG TTC CGC CTG AAA CAC AGA CCC TTC CAC GTG GTG CTC AGC ATG TTC CTC AGC CTTG TAT CTC AGG AGG CAC CTC GTG CTC AGC ATG TTC CTC AGC CTTG TAT CTC AGG AGG CAC CTC GTG CTC AGC CTC TAC ATG CTA CTC CTG TAT CTC AGG	GTCGACTCTA GAGTGTGTGT CAGCACTTGG CTGGGGACTT CTTGAACTTG 60 70 80 90 100 CAGGGAGAAT AACTTGCGCA CCCCACTTG CGCCGGTGCC TTTGCCCCAC 110 120 130 140 150 CGGAGCCTGC TTCGCCATCT CCGAGCCCCA CCGCCCCTCC ACTCCTCGGC 160 170 180 190 200 CTTGCCCGAC ACTGAGACGC TGTTCCCAGC GTGAAAAGAG AGACTGCGCC 210 220 230 240 250 GCCGGCACCC GGGAGAAGGA GGAGGCAAAG AAAAGGAACG GACATTCGGG 260 270 280 290 300 CCTTGCGCCA GGTCCTTTGA CCAGAGTTTT TCCATGTGGA CGCTCTTTCC 310 320 330 340 350 ATGGACGTGT CCCCGCGTGC TTCTTAGACG GACTGCGGTC TCCTAAAGGG (1) 370 385 40 CGACC ATG GTG GCC GGG ACC CGC TGT CTT CTA GCG TTG CTG CTT CCC CTMET Val Ala Gly Thr Arg Cys Leu Leu Ala Leu Leu Leu Pro Gl CTC CTG GCC GCC GCC GCC CTC GTT CCC GAG CTG GCC CGC AGG AAG Leu Leu Gly Gly Ala Ala Gly Leu Val Pro Glu Leu Gly Arg Arg Lys 460 475 GCG GCG TCG TCG CCC CCC TCA TCC CAG CCC TCT CCC CTC ALA ALA Ser Ser Gly Arg Pro Ser Ser Gln Pro Ser Asp Glu Val Leu TTC GAG TTG CGC CTG CTC ACC ATG TTC CGC CTG AAA CAG AGA CCC ACC TTC GAG TTG CGC CTG CTC ACC ATG TTC CGC CTG AAA CAG AGA CCC ACC TTC GAG TTG CGC CTG CTC ACC ATG TTC CGC CTG AAA CAG AGA CCC ACC TTC GAG TTG CGC CTG CTC ACC ATG TTC CGC CTG AAA CAG AGA CCC ACC TTC GAG TTG CGC CTG CTC ACC ATG TTC CGC CTG AAA CAG AGA CCC ACC TTC GAG TTG CGC CTG CTC ACC ATG TTC CGC CTG AAA CAG AGA CCC ACC TTC GAG TTG CGC CTG CTC ACC ATG TTC CGC CTG AAA CAG AGA CCC ACC TTC GAG TTG CGC CTG CTC ACC ATG TTC CGC CTG AAA CAG AGA CCC ACC TTC GAG TTG CGC CTG CTC ACC ATG TTC CGC CTG AAA CAG AGA CCC ACC TTC GAG TTG CGC CTG CTC ACC ATG TTC CGC CTG AAA CAG AGA CCC ACC TTC GAG TTG CGC CTG CTC ACC ATG TTC CGC CTG AAA CAG AGA CCC ACC TTC GAG TTG CGC CTG CTC ACC ATG TTC CGC CTG AAA CAG ACA CCC ACC TTC GAG CGC CTG CTG CTC ACC ATG TTC CGC CTG CTG CTG CTG CTG CTG CTG C	GTCGACTCTA GAGTGTGTGT CAGCACTTGG CTGGGGACTT CTTGAACTTG 60 70 80 90 100 CAGGGAGAAT AACTTGCGCA CCCCACTTTG CGCCGGTGCC TTTGCCCCAG 110 120 130 140 150 CGGAGCCTGC TTCGCCATCT CCGAGCCCCA CCGCCCCTCC ACTCCTCGGC 160 170 180 190 200 CTTGCCCGAC ACTGAGACGC TGTTCCCAGC GTGAAAAGAG AGACTGCGCG 210 220 230 240 250 GCCGGCACCC GGGAGAAGGA GGAGGCAAAG AAAAGGAACG GACATTCGGT 260 270 280 290 300 CCTTGCCCCA GGTCCTTTGA CCAGAGTTTT TCCATGTGGA CGCTCTTTCA 310 320 330 340 350 ATGGACGTGT CCCCGCGTGC TTCTTAGACG GACTGCGGTC TCCTAAAGGT (1) 370 385 400 GAACC ATG GIG GCC GGG ACC GCC TGT CIT CIA GCC TIG CIG CIT CCC CAG GC MET VAI Ala Gly Thr Arg Cys Leu Leu Ala Leu Leu Pro Gln Vi 415 430 445 CTC CTG GCC GCC GCG GCT GCC CTC GTT CCG GAG CTC GCC ACG ACG AAG TTC Leu Leu Gly Gly Ala Ala Gly Leu Val Pro Glu Leu Gly Arg Arg Lys Phe 460 475 GCG GCG TCG TCG GCC CCC TCA TCC CAG CCC TCT GAC GAG GTC CTG AGC 120 120 120 120 120 120 120 120 120 120

		625		563			640					655					670		
	Gln	Pro	GGC Gly	Ser		GCC	CCA Down	GAC	CAC	∆~~	TIG	GAG	AGG	GCA Ala	GCC	AGC	ΩGA λ	GCC N	
5				-	110	Λια	PIU	קבת	בינו	мg	Leu	GIG	wa	via	ALG	Ser	Mrg	ALG	
				685					700					715					
	AAC	ACT	GIG	œc	AGC	TTC	CAC	CAT	GAA	CAA	TCT	TTG	GAA	GAA	ĞIY.	CCA	GAA	ACG	
	NO:1	1115	Val	Arg	Ser	Pne	HIS	Fus	GIū	GIU	Ser	Leu	GIU	GIU	Leu	Pro	GIU	'Inr	
10											760					775			
	730	ccc	AAA	302	λCC	745)C)	حكس	THY :	ململمة	760		AGT	TCT	ATC	775 ex	ACG	GAG	
	Ser	Gly	Lys	Thr	Thr	Arg	Arg	Phe	Phe	Phe	Asn	Leu	Ser	Ser	Ile	Pro	Thr	Glu	
		_				_	_									-			
15	CAG	ململت	790 ATC	200	m-a	CCA	CAG	805	CAC.	ملتك	مكنت	CC3	820 GAA	CAG	ATG	CAA	GAT	835 GCT	
	Glu	Phe	lle	Thr	Ser	Ala	Glu	Leu	Gln	Val	Phe	Arg	Glu	Gln	MET	Gln	Asp	Ala	
																	_		
	OTTIN	CC3	AAC		850	3.000		~~		865		220	N/TTP	ጥእጥ	880	MY	ATTA	222	
20	Leu	Glv	Asn	ASD	Ser	Ser	Phe	His	His	Aru	Ile	ASTI	Ile	Tyr	Glu	Πe	Πe	Lys	
20		-																_	
	~~	895					910		~ ~	~~~		925	- CTITI	ann-	CNC	3~	940	ante.	
	Pro	Ala	Thr	Ala	ASO	Ser	AAA Lws	Phe	Pro	Val	Thr	Ser	Leu	Leu	Asp	Thr	Arq	TIG Leu	
25 `	•••				•	-	_1_								•				
25				955				-	970				~~~	985		com	~TT	אחר	
																		ATG MET	
•	-																		
•	100					1015					1030			~~~		1045		con co	
30	Ares	Tro	ACI	GCA Ala	CAG	GGA	CAC Hie	Ala	AAC	His	G	Phe	. GIG	Val	Glu	Val	Ala	CAC	
						75 G.		W ,7					જ જે	. ; <,			• •		•
			1060					1075					1090					1105	
																		TIG Leu	
35					-	. 027					,							-	
					1120					1135					1150		-	. ~~	
																		GCC	
						,						,						4	
40		1165					1180					119					1210		
	His	AST	GUA	L AA	GGC	CAT His	. CC1	CIC	CAC 1 His	: AAJ	A AG	A GAV	L AAA	. Arr	r Glr	A GCC	INS	CAC His	
	•		, (1)		- 023	•	,			<i>.</i>		,			,				
				1225					1240					125				. ~~	
45																		GAC Asp	
	,			, -,,	, 44	، حد	ı ıy.	,		. .	، رب		۔۔. و			1-		<u>-</u> -	
	12					128					130					131			_
	TT	C AG	T GA	C GI	G GG	g TG	G AA	I GA	ic Te	G AI	T GI	G GC	TCC	$x = \infty$	2 GG	GTA	r Ca - Wi	C GCC) A
50	Fil	e 5e	133		T 61	УТТ	p ws	11 AS		בי ק	e vo	יי אי	.a Pi 136		O GI	y iy.		s Al a 137	
	TI	T			c ca	iA Gi	A TO			m «	II C	rg G			a c	eg aa	c T	C YC	
																		ar IV	
					139	20				140	าร				142	20			
55	AA	T C	T GC	e ai			S AC	I I	rs G			cr G	T A	ic T			πα	T AA	G
	<u> </u>	n Hi	s Al	a II	e Va	n c	n T	r D	en W	<u>م</u> ا له	sn S	er V	al A	en Se	er ly	ys II	e P	n Ly	S

1435 1450 1465 1480 GCA TGC TGT GTC COG ACA GAA CTC AGT GCT ATC TCG ATG CTG TAC CTT GAC GAG Ala Cys Cys Val Pro Thr Glu Leu Ser Ala Ile Ser MET Leu Tyr Leu Asp Glu 5 1510 1525 AAT GAA AAG GIT GIA TIA AAG AAC TAT CAG GAC ATG GIT GIG GAG GGT TGT GGG Asn Glu Lys Val Val Ieu Lys Asn Tyr Gln Asp MET Val Val Glu Gly Cys Gly 10 1540 (396) 1553 1563 1573 1583 1593 1603 TGT CGC TAGTACAGCA AAATTAAATA CATAAATATA TATATATATA TATATTTTTAG AAAAAAGAAA Cys Arg 15 AAAA, wobei das Verfahren die nachfolgenden Schritte umfaßt: (a) Absuchen einer Genbank durch Hybridisieren mit einem markierten bBMP-2-Fragment, wobei die Genbank 20 aus einer von U-2 OS abgeleiteten DNA oder cDNA konstruiert war, (b) Isolieren positiver Clone und (c) Isolieren der DNA-Insertionen aus diesen Clonen. 25 Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß das Gen menschliches BMP-2 codiert, das die in Anspruch 1 angegebene Aminosäuresequenz aufweist. Verfahren zur Herstellung eines Gens, das ein Protein codiert, das Eigenschaften von menschlichem BMP-2 zeigt, 30 und eine DNA-Sequenz umfaßt, die: (a) sich in der Codonsequenz infolge der Degeneriertheit des genetischen Codes von einer DNA-Sequenz nach Anspruch 1 unterscheidet; (b) mit einer DNA-Sequenz nach Anspruch 1 oder nach vorstehendem Absatz (a) hybridisiert; oder 35 (c) ein Fragment, eine allelische oder eine andere Variation einer DNA-Sequenz nach Anspruch 1 darstellt, unabhängig davon, ob die Variation zu Änderungen in der Peptidsequenz führt oder nicht, 40 wobei das Verfahren Standardtechniken der Molekularbiologie umfaßt. 4. Verfahren nach Anspruch 3, dadurch gekennzeichnet, daß die DNA-Sequenz eine genomische DNA-Sequenz ist. 5. Verfahren nach Anspruch 3, dadurch gekennzeichnet, daß die DNA-Sequenz eine cDNA-Sequenz ist. 45 6. Verfahren zur Herstellung eines Rinder-BMP-2 codierenden Gens, umfassend die nachfolgende DNA-Sequenz: 50

```
30
                       15
       GGC CAC GAT GGG AAA GGA CAC CCT CTC CAC AGA AGA GAA AAG CGG
                      X
                          G
                                             3
                                                              90
                       60
       CAA GCA AAA CAC AAA CAG CGG AAA CGC CTC AAG TCC AGC TGT AAG
                              3
10
                                          120
       AGA CAC CCT TTA TAT GTG GAC TTC AGT GAT GTG GGG TGG AAT GAC
                              D
                                          D
                                          155
                      150
15
       TGG ATC GTT GCA CCG CCG GGG TAT CAT GCC TTT TAC TGC CAT GGG
                          P
                              G
                                  Y
                                                             225
                      195
                                         . 210
       GAG TGC CCT TTT CCC CTG GCC GAT CAC CTT AAC TCC ACG AAT CAT
20
           C
               P
                                  D
                                                 S
                          L
                              Α
                                      H
                                         L
                                             N
                  F
                                          255
                      240
       GCC ATT CTC CAA ACT CTG GTC AAC TCA GTT AAC TCT AAG ATT
                          L
                              V
                                  N
                                      S
                                             N
25
                             300
              38.5
       AAG GCA TGC TGT GTC CCA ACA GAG CTC AGC GCC ATC TCC ATG CTG
                          5
                              T
                                  Ξ
                                          S
30
                       330
                                          345
        TAC CTT GAT GAG AAT GAG AAG GTG GTA TTA AAG AAC TAT CAG GAC
                      N
                                  V
                                      V
                                              X
                       375 -
                                       (129)
                                                 397
35
        Ξ
                               G
                                   С
                                      \mathbb{R}
                       G
                           С
40
                                               447
                                                          457
               417
                         427
                                    437
        TAAATATATA TATATATAT TTAGAAAAAC AGCAAAAAAA TCAAGTTGAC
                                    487
                         477
        ACTITAATAT TTCCCAATGA AGACTITATT TATGGAATGG AATGGAGAAA
45
                                                          557
               517
                         527
                                    537
                                               547
        AAGAAAACA CAGCTATTTT GAAAACTATA TTTATATCTA CCGAAAAGAA
                         577
                                    587
50
        GTTGGGAAAA CAAATATTTT AATCAGAGAA TTATT
```

wobei das Verfahren die nachfolgenden Schritte umfaßt:

- (a) Absuchen einer Genbank mit einer markierten auf der Grundlage der Aminosäuresequenz eines Fragmentes von bBMP-2 entworfenen Sonde, wobei die Genbank aus Rinderleber-DNA oder cDNA konstruiert wurde,
 - (b) Isolieren positiver Clone und

- (c) Isolieren der DNA-Insertionen aus diesen Clonen.
- 7. Verfahren nach Anspruch 6, dadurch gekennzeichnet, daß das Gen Rinder-BMP-2 codiert, das die Aminosäuresequenz von Anspruch 6 aufweist.
- 8. Verfahren zur Herstellung eines Genes, das ein Protein codiert, das Eigenschaften von Rinder-BMP-2 zeigt, und DNA-Sequenzen umfaßt, die:
 - (a) sich in der Codonsequenz infolge der Degeneriertheit des genetischen Codes von einer DNA-Sequenz nach Anspruch 7 unterscheiden;
 - (b) mit einer DNA-Sequenz nach Anspruch 7 oder nach vorstehendem Absatz (a) hybridisieren; oder
- (c) Fragmente, allelische oder andere Variationen einer DNA-Sequenz nach Anspruch 7 darstellen, unabhängig davon, ob die Variationen zu Änderungen in der Peptidsequenz führen oder nicht,

wobei das Verfahren Standardtechniken der Molekularbiologie umfaßt.

- 9. Verfahren nach Anspruch 8, dadurch gekennzeichnet, daß die DNA-Sequenz eine genomische DNA-Sequenz ist.
- 10. Verfahren nach Anspruch 8, dadurch gekennzeichnet, daß die DNA-Sequenz eine cDNA-Sequenz ist.
- 11. Verfahren zur Herstellung eines menschliches BMP-4 codierenden Genes, das die nachfolgende DNA-Sequenz umfaßt:

	4.1.1.4.1				
25	10 CTCTAGAGGG	20 CAGAGGAGGA	30 GGGAGGGAGG	40 GAAGGAGCGC	50 GGAGCCCGGC
30	60	70 GGTGAGTGTG	80 GCATCCGAGC	90 TGAGGGACGC	100 GAGCCTGAGA
, 55	110	120 GCTCCGGCTG	130	140	150
35	160	170	180	190	200 GCCCTCGCCC
	210	220	230	240	
40	260	270	280	290	300
	. 310	320	330	340	
45	360	370	380	390	
	CTGTCAAGAA	TCATGGACTO	TTATTATAT	CCTTGTTTTC	TGTCAAGACA

50

5

10

15

20

(1) CC ATG ATT CCT MET Ile Pro

-																			
			c										447 CAA Gln						
10				CT .									YYY YYY						
15	522 GGC Gly	CA(C G	cs (ejà ecy	gĵλ œΥ	537 CGC Arg	CGC Arg	TCA Ser	età eee	CAG Gln	552 AGC Ser	CAT His	GAG Glu	CTC Leu	CIG Leu	567 CGG ਮੁੱਤ	yzb GYC	TTC Phe
20	GAG Glu	GC Al	S	582 ACA Err	CIT Leu	CTG Leu	CAG Gln	ATG MET	597 TIT Phe	egy eee	CIG Leu	CSC Arg	Arg CGC	612 CSC Arg	CCG Pro	CAG Gln	CCT Pro	AGC Ser	627 AAG Lys
25	AGT Ser	GC Al	æ (STC Val	ATT Ile	642 CCG Pro	GAC Asp	TAC Tyr	atg Met	ಜ್ ಜ್	657 දූදු දූදු	CTT Leu	TAC Tyr	CGG Arg	CIT Leu	672 CAG Gln	TCI Ser	G G G G	G]u
	GAG Glu	68 GA G1	S (IAA Ilu	GAG Glu	CAG Gln	ATC Ile	702 CAC His	AGC Ser	ACT	GT Gly	CTT Leu	717 GAG Glu	TAT Tyr	CCI Pro	GAG Glu	% €	732	c∞ Ala
30						: Aα					CAC					CTG			: ATC
35		G						TCI					CIC					AGC	ATC : Lle
40					GA(CC					TIC				897 GIG Val
							rac					CAC					נגע		GTT Val
45		G A							GIX					OIA C					e cac Asp
50						C C3(λŒ					r GAI			: 3rp : ccr
5 5		S G											A AAG					_ AT	r GAG e Glu

	1122 1157 1152 1167
	CLE YOU CAN CLO CAL CAN YOU CON YOU CAN CON CAN CAN CLO YER WILL ARC
,	Val Thr His Leu His Gln Thr Arg Thr His Gln Gly Gln His Val Arg Ile Ser
5	
-	1182 1197 1212
	CCA TOC THA OUT CAA CCC ACT CCC AAT TCC CCC CAC CTC CCC CC
	Ary Ser Leu Pro Glin Gly Ser Gly Asn Try Ala Glin Leu Ary Pro Leu Leu Val
	1997 1940 1959 1979
10	1227 1242 1257 1272
	ACC TIT GGC CAT GAT GGC GGG GGC CAT GCC TIG ACC CGA GGC GGG AAG
	The Gly His Asp Gly Arg Gly His Ala Leu The Arg Arg Arg Arg Ala Lys
	1287 1302 1317
	CGT AGC CCT AAG CAT CAC TCA CAG CGG GCC AGG AAG AAG AAT AAG AAC TGC CGG
15	Arg Ser Pro Lys His His Ser Gln Arg Ala Arg Lys Lys Asn Lys Asn Cys Arg
	3 - 2 - 2 - 2 - 2 - 2 - 2 - 2 - 2 - 2 -
	1332 1347 1362 1377
	CCC CAC TOS CTC TAT GTG GAC TTC AGC GAT GTG GGC TGG AAT GAC TGG ATT GTG
20	Ary His Ser Leu Tyr Val Asp The Ser Asp Val Gly Trp Asn Asp Trp Ile Val
20	and the set led the Am we like set tob ter out out out out out
	1392 1407 1422 1437
	GCC CCA CCA GGC TAC CAG GCC TTC TAC TGC CAT GGG GAC TGC CCC TTT CCA CTG
	Ala Pro Pro Gly Tvr Gln Ala Phe Tvr Cvs His Glv Aso Cvs Pro Phe Pro Leu
25	2402
	GCT GAC CAC CTC AAC TCA ACC AAC CAT GCC ATT GTG CAG ACC CTG GTC AAT TCT Ala Asp His Leu Asn Ser Thr Asn His Ala Ile Val Gln Thr Leu Val Asn Ser
	Ala Asp his leu Ash Ser thir Ash his Ala Tie Val Gill his 220 to 120 to
	1497 1512 1527 1542
30	CITY AND THE ACT ATE OF AAA COU TET TET GTG COO ACT GAA CTG AGT COO ATE
	Val Asn Ser Ser Ile Pro Lys Ala Cys Cys Val Pro Thr Glu Leu Ser Ala Ile
	1557 1572 1587
35	TOO ATG CTG TAC CTG GAT GAG TAT GAT AAG GTG GTA CTG AAA AAT TAT CAG GAG
00	Ser MET Leu Tyr Leu Asp Glu Tyr Asp Lys Val Val Leu Lys Asn Tyr Gln Glu
	3656
	ATG GTA GTA GAG GGA TGT GGG TGC GGC TGAGATCAGG CAGTCCTTGA GGATAGACAG
40	MET Val Val Glu Gly Cys Gly Cys Arg
45	
	1666 1676 1686 1696 1706
	ATATACACAC CACACACACA CACCACATAC ACCACACACA
50	
	1716 1726 1736 1746 1756
	TCCACTCACC CACACACTAC ACAGACTGCT TCCTTATAGC TGGACTTTTA
	1766 1776 1786 1796 1806
55	TTTAAAAAA AAAAAAAAA AATGGAAAAA ATCCCTAAAC ATTCACCTTG

ACCTTATTTA TGACTTTACG TGCAAATGTT TTGACCATAT TGATCATATA

1826

1816

1836

1846

1856

5	1866 1876 1886 1896 1906 TTTTGACAAA ATATATTTAT AACTACGTAT TAAAAGAAAA AAATAAAATG
10	1916 1926 1936 1946 AGTCATTATT TTAAAAAAAA AAAAAAAACT CTAGAGTCGA CGGAATTC
	wobei das Verfahren die nachfolgenden Schritte umfaßt:
15	(a) Absuchen einer Genbank durch Hybridisieren mit einem markierten bBMP-2-Fragment, wobei die Genbank aus einer von U-2 OS abgeleiteten DNA oder cDNA konstruiert war,
	(b) Isolieren positiver Clone und
	(c) Isolieren der DNA-Insertionen aus diesen Clonen.
20	 Verfahren nach Anspruch 11, dadurch gekennzeichnet, daß das Gen menschliches BMP-4 codiert, das die in Anspruch 11 angegebene Aminosäuresequenz aufweist.
25	13. Verfahren zur Herstellung eines Genes, das ein Protein codiert, das Eigenschaften von BMP-4 zeigt, und eine DNA-Sequenz umfaßt, die:
-	(a) sich in der Codonsequenz infolge der Degeneriertheit des genetischen Codes von einer DNA-Sequenz nach Anspruch 11 unterscheidet;
30	(b) mit einer DNA-Sequenz nach Anspruch 11 oder vorstehendem Absatz (a) hybridisiert; oder
55	(c) ein Fragment, eine allelische oder eine andere Variation einer DNA-Sequenz nach Anspruch 11 darstellt, unabhängig davon, ob die Variation zu Änderungen in der Peptidsequenz führt oder nicht,
35	wobei das Verfahren Standardtechniken der Molekularbiologie umfaßt.
00	14. Verfahren nach Anspruch 13, dadurch gekennzeichnet, daß die DNA-Sequenz eine genomische DNA-Sequenz ist.
	15. Verfahren nach Anspruch 13, dadurch gekennzeichnet, daß die DNA-Sequenz eine cDNA-Sequenz ist.
40	16. Vektor, enthaltend das Gen oder die DNA-Sequenz nach einem der Ansprüche 1 bis 15 in einer funktionellen Verbindung mit einer Expressions-Kontrollsequenz.

- 17. Zelle, dadurch gekennzeichnet, daß sie mit einem Vektor nach Anspruch 16 transformiert ist.
- 45 18. Zelle nach Anspruch 17, dadurch gekennzeichnet, daß sie eine Säugerzelle, eine Bakterienzelle, eine Insektenzelle oder eine Hefezelle ist.
 - 19. Zelle nach Anspruch 18, dadurch gekennzeichnet, daß sie eine CHO-Zelle ist.
- 50 20. Verfahren zur Herstellung eines Proteins, das Eigenschaften von BMP-2 zeigt, umfassend die Schritte
 - Züchten einer mit einem Expressionsvektor transformierten Zelle in einem geeigneten Kulturmedium, wobei der Expressionsvektor ein Gen oder eine DNA-Sequenz umfaßt, die nach einem der Ansprüche 1 bis 10 hergestellt wurden, und
 - Gewinnen des Proteins aus dem Kulturmedium.

55

21. Verfahren zur Herstellung eines Proteins, das Eigenschaften von BMP-4 zeigt, umfassend die Schritte

- Züchten einer mit einem Expressionsvektor transformierten Zelle in einem geeigneten Kulturmedium, wobei der Expressionsvektor ein Gen oder eine DNA-Sequenz umfaßt, die nach einem der Ansprüche 11 bis 15 hergestellt wurden, und
- Gewinnen des Proteins aus dem Kulturmedium.
 - 22. Verfahren zur Herstellung eines Proteins, das Eigenschaften von BMP-2 oder BMP-4 zeigt, umfassend die Schritte
 - Züchten der Zelle nach Anspruch 17 in einem geeigneten Kulturmedium und
 - Isolieren des Proteins aus dem Kulturmedium.
 - 23. Verfahren zur Herstellung eines Arzneimittels, dadurch gekennzeichnet, daß es ein Kombinieren der nach einem der Ansprüche 20 bis 22 hergestellten Proteine, einzeln oder in Kombination, mit einem pharmakologisch verträglichen Träger umfaßt.
 - 24. Verfahren nach Anspruch 23, dadurch gekennzeichnet, daß das Arzneimittel ferner eine Matrix umfaßt, die fähig ist, das Arzneimittel an die Stelle des Knochen- oder Knorpelschadens zu liefern und eine Struktur zur Induktion der Knochen- oder Knorpelbildung bereitzustellen.
 - 25. Verfahren nach Anspruch 24, dadurch gekennzeichnet, daß die Matrix Hydroxyapatit, Kollagen, Polyessigsäure oder Tricalciúmphosphat umfaßt.
- 26. Verwendung eines Proteins nach einem der Ansprüche 20 bis 22, einzeln oder in Kombination, zur Herstellung eines Arzneimittels zur Induktion der Knochen- oder Knorpelbildung.

Revendications

10

15

30

35

40

45

50

55

Revendications pour les Etats contractants suivants : BE, CH, DE, FR, GB, IT, LI, LU, NL, SE

1. Gène codant pour la BMP-2 humaine comprenant la séquence d'ADN suivante :

		10			20		3	0		.40			50			60		70
	GICCAC	ICIA	CYCI	GIGI	GI (yecy	CIIG	G CI	CCCC	ACIT	CII	CYYC	TTG	CAGG	CYCY.	AT 2	AACIT	CCCC
5																		
		80			90	•	10	n		110			120		٠ ۾	30	.•	140
	com Ac		α			TTGC		-	SAGO					ಯಾಸಿ			∞	
10				_											_			
	سخست	150	~		60 20 }	~~··	17			180			190			00 ~~	~~~	210
	ACTOOT	سعد	CIIG	شللك	ه علم	نحالانا	مانحا	C 16	1114	ساحلامه	خال	بالالماما	العطالع	سالمشكلت	يلنن	ا ول	حللك	سالن
15		220		2	30		24	0		250			260		2	70 [†]		280
	GGGAGA	agga	CCAC	GCAA	yc y	<i>AAA</i> G	GAA0	G GA	CATT	ŒŒ	α	TGC	CCA	GGTO	CITR	EA C	ككوكز	GITT'
		290		3	۵0،		.31	^		320			330		2.	40		:350
	TCCATG	-	ŒŒ			TGCA								GACT			גיוסטי	
20																		
		(<u>1)</u>	~~ ~			70 ~~ ~		<u>.</u> .			55 					00		
	CEACE	KET V																
25	•		- A	C.	-y -	· ^-	-9 C	ys D		בנו אנ	וב ט			eu r.	. O G.	v	شت	
			415					430					445					
	arc arc	G GGC	: GGC	ccc	GCT	GGC	CIC	CIT	α	ದಿಸಿದ	CIG	GGC	œc	AGG	770	TTC	: ccc	
	Leu Le	ı CīĀ	Gly	Ala	Yjs	GŢĀ	Ieu	Val	Pro	Glu	Leu	GŢĀ	Arg	Arg	īvs	Phe	: Ala	
30	450				475					490					505			
	<u>ೀ</u> ಆದ	i Tai	TOG				TCA	TCC	CNG		TOT	GAC	೦೭೦	GTC		250	GAG:	
	Ala Ala																	
	·										•							
35	TIC GA	520 520		بكتك	سبب	100	535			~~~		550	363	~~	•	~~	565	
35	Phe Glu	ı Leu	محی ،	Ten	Ten	حنہ. حنہ	rc	The	Glv	Tan	TVE	C40	للتلا	2-7	4	مبرد	Ser	
			– 9						<u></u>		د پرد							
				580					595					510				
	AGG GAG																	
40	भूति भूद्र	בבה כ	. Val	۱ع۷	520	520	īĀī	MET	Leu	خحز	ŢĠIJ	TÀT	لمتر	وسد	عنة:	Sex	. CTĀ	
	625	5				640					635					570)	
	ಯ ಜಾ		TCA	∞	GΩ		GAC	Cý.C	ಯ	TIG		AGG	GCA	GΩ	AGC			
	Gln Pro																	
45							-					3						
	\\C \~	n ~	685					700					715				.~~	
	AAC AC	- 1127 T. CTC	. ಹಾರ - ಹಾರ	AGC	TIC	CXC	CAI	GAA	C7Y	TCT	TIG	G5A	CAA	CIA	2	دینے ۱۰۰ء	هنکاتر ا	
	Asa Thi	٠ ٧٥١	अपु	Ser	7.21€	HIZ	عد:	CT.II	CTIT	ser	ren	Giù	GIU	Leu	220	نيد		

	730					745					760					775		
	ACT	GGG	AAA	ACA	ACC	α G	AGA	TTC	TTC	TTT	TAA	TIA	AGT	TCT	ATC	∞	ACG	GYG
5	Ser	Gly	Lys	Thr	Thr	Arg	Arg	Phe	Phe	Phe	Asn	Leu	Ser	Ser	Ile	220	The	Glu
			790			•		805					820					835
	GAG	TTT		∞ A	TCA	GCA	GAG		CAG	GTT	TTC	ŒÀ		CAG	λIG	CAA	CAT	CCT
																		Ala
10																		
10					850					865					880			
																		AAA
	ieu	СīĀ	ASD	ASN	Ser	Ser	Pne	HIS	HIS	Arg	Пе	ASD	116	TÄL	GLU	тте	TTE	Lys
		895					910					925					940	
15	CCT		ACA.	GC	220	TY		علىك	~	CIC	ACC		CITT	TIG	GAC	AΩ		TTG
	Pro	Ala	Thr	Ala	Asn	Ser	INS	Phe	Pro	Val	Thr	Ser	Leu	Leu	ASO	Thr	Arq	Leu
															•		_	
				955					970					985				
20	GIG	AAT	CAG	AAT	GCA	AGC	AGG	TGG	CFY	AGT	TIT	GAT	GIC	ACC	∞	GCT	GIG	ATG
	Val	ASN	Gln	YZU	Ala	Ser	Arg	đĩ	Glu	Ser	Phe	yzō	Val	Thr	Pro	Ala	Val	MET
	100	•													-	045		
	1000					1015					1030	mm~	~~~	~~~		L045	~~	Cic
	2	166	ACT.	GCA 33.5	ملات	COA	CAC	312	المند	CAT.	Clar	THO	CIG	GIG	Clu	Maj.	Ala	CAC His
25	وجم	112	111L.	MIA	GIII	GTĀ.	מנת	ALC	~211	ri2	GTÅ	File	٧ حــــــــــــــــــــــــــــــــــــ	بنيتا	عادى	٧	7	-
			1060				-	1075					1090					1105
	TTG			AAA	CAA	.GGT	GIC	$\tau \infty$	AAG	AGA	CAT			ATA	AGC	AGG	TCT	TTG
																		ren
30																		
					1120					1135					1150			
																		GGC
	als	GTJ	YZD	GIU	Fis	Ser	JID	Ser	GTIJ	TTE	Arg	Pro	Leu	Leu	٧ڪي	THE	-4:6	Gly
		1155			•		1180					1195					1210	
35			GGA	AAA	GGG			CTC	CAC	AÁA			AAA	œ	СУУ	-		CAC
	His	ASD	Gly	Lys	Glv	His	Pro	Leu	His	Lvs	Arq	Glu	Lvs	Arg	Gln	Ala	Lys	His
		-	-	-						-	_		-				-	
			-	1225					1240					1255				
40	AAA	೧೫೦	α	AAA	α C	CIT	AAG	TCC	AGC	TGT	AAC	AGA	CAC	CCT	TIG	TAC	GIG	GYC
	Lys	Gln	Arg	ΓÃ2	ਮੁਸ਼ਕੇ	Leu	Lys	Ser	Ser	Cys	Lys	Arg	Fis	STO	Teu	Tyr	VZI	yzb
	1270	,				205					1200				-	1315		
			cic	~		1285	2.20	C1C			T300	ىلىت	~~	يىبى			CiC	ထေ
45	270	252	300	UzU UzU	Clv	₩	7 cm	350	مجي ا	TIO	Val	212	2	D	Glv	40	His	Ala
40		-	ມລຸ	V 223,	GTĀ	בבב		تيم	יבב	110	V							
		3	1330				-	1345					1350				:	1375
	$\overline{\text{TT}}$	TAC	TGC	CAC	GGA	GAA	TGC	α	TIT	α	CIG	CCT	G-T	CAIT	CIG	AAC	TCC	ACE
	Phe	$\mathbb{T}_{2^{n}}$	Cys	His	Gly	Glu	Cys	Pro	Phe	ನ್ನು	ĭeu	Ala	نحز	<u>Hi</u> s	Leu	מצג	Ser	Thi
50					_													
			_		L390					1405					1420		~~	
	AAT	GT.	$\epsilon \infty$	ATT	CIT	CyC	ΑŒ	TIG	CIC.	AAC	TCT	.GII	AAC	TCT	عدي	ATT	CI	AAG
	:Sn	His	Ala	Ile	Val	Gln	Thr	Leu	٧ <u>حا</u>	لتحتر	Ser	Val	ASN	Ser	Lys	Ile	520	_7 75

		1435 1450 1465 1480
		eca tec tet etc coe aca eaa ctc aet ect atc toe ate cte tac ctt eac eae
5		Ala Cys Cys Val Pro Thr Glu Leu Ser Ala Ile Ser MET Leu Tyr Leu Asp Glu
•	•	1495 1510 1525
		AAT GAA AAG GIT GIA TITA AAG AAC TAT CAG GAC ATG GIT GIG GAG GGT TGT GGG
		Asn Glu Lys Val Val Leu Lys Asn Tyr Gln Asp MET Val Val Glu Gly Cys Gly
10		1540(396) 1553 1563 1573 158 3 1593 1 603
, ,		1540 (396) 1553 1563 1573 1583 1593 1603 TGT CGC TAGTACAGCA AAATTAAATA CATAAATATA TATATATATA TATATTITAG AAAAAAAAA
		Cys Arg
15		AAAA
	2.	Gène codant pour la BMP-2 humaine comportant la séquence d'acides aminés donnée à la revendication 1.
	2	Gène codant pour une protéine montrant des propriétés de la BMP-2 humaine et comprenant une séquence
20	J .	d'ADN :
		(a) qui diffère d'une séquence d'ADN de la revendication 1 dans la séquence de codons du fait de la dégé-
		nérescence du code génétique ; (b) qui s'hybride avec une séquence d'ADN de la revendication 1 ou du paragraphe (a) ci-dessus ; ou
25		(c) représente un fragment, une variation allélique ou autre d'une séquence d'ADN de la revendication 1, que
		cette variation résulte de changements dans la séquence peptidique ou non.
	4	Séquence d'ADN quivent le revendication C qui est une séquence d'ADN cénemique
	4.	Séquence d'ADN suivant la revendication 3, qui est une séquence d'ADN génomique.
30	5.	Séquence d'ADN suivant la revendication 3, qui est une séquence d'ADNc.
		O) and a second DMD O1 at a second of the se
	6.	Gène codant pour la BMP-2 bovine comprenant la séquence d'ADN suivante :
35		
40		
		•
45		
		•

5	(1) GGC G	CAC	GAT D	GGG G	15 AAA K	GGA	CAC H	CCT P	CTC L	30 CAC H	AGA R	AGA R	GAA E	AAG K	45 CGG R
10	CAA Q	GCA A	AAA K	CAC H	60 AAA K	CAG Q	CGG R	AAA K	CGC R	75 CTC L	AAG K	TCC S	AGC S		
	AGA R	CAC H	CCT P	TTA L	105 TAT Y	GTG V	GAC D	TTC F	AGT S	120 GAT D	GTG V	GGG G	TGG W	AAT N	135 GAC D
15	TGG W	ATC I	GTT V	GCA A	150 CCG P	CCG P	GGG G	TAT Y	CAT H	165 GCC A	TTT F	TAC Y	TGC C	CAT H	180 GGG G
20	GAG E	TGC	CCT P	TTT F	195 CCC P	CTG L	GCC A	GAT D	CAC H	210 CTT L	AAC N	TCC S	ACG T	AAT N	225 CAT H
25	GCC A	ATT	V CTC	CAA Q	240 ACT T	CTG: L	GTC V	AAC N	TCA S	255 GTT V	AAC	TCT	'AAG K	ATT I	270 CCC P
30	AAG	GCA	TGC	TGT	GTC	CCA	ACA	GAG	CTC L	AGC	GCC	ATC	TCC	3 ጥር	315 CTG L
	TAC Y	CTT L	GAT D	GAG E	330 AAT N	GAG E	AAG K	GTG V	GTA V	345 TTA L	AAG K	AAC N	TAT Y	CAG O	360 GAC D
35	ATG M	GTT V	GTC V	GAG E	GGT	TGT	GGG	TGT	(129 CGT R	TAGO	3 ACAC	97 CA A	AATA	40 LAAAT	7 CA
40	TAA	4 ATATA	17 TA T	'ATAT	42 ATAT	7 A TT	'AGAA	437 AAAC	AGC	AAAA	447 AAA	тсла	4 GTTG	57 AC	
	ACTI		67 AT T	TCCC	47 AATG	7 A AG	ACTT	487 TATT	TAT	'GGAA	497 TGG	AATG	5 GAGA	07 AA	
45		5	17		52	7		537			547		=	= 7	
	AAGA	AAAA		AGCT			AAAC	TATA	TIT	'ATAT	CTA	CCGA	AAAG	AA	

- 7. Gène codant pour la BMP-2 bovine contenant la séquence d'acides aminés de la revendication 6.
- 8. Gène codant pour une protéine montrant des propriétés de la BMP-2 bovine et comprenant des séquences d'ADN :

(a) qui diffèrent d'une séquence d'ADN de la revendication 7 dans la séquence des codons du fait de la dégénérescence du code génétique ;

- (b) qui s'hybrident avec une séquence d'ADN de la revendication 7 ou du paragraphe (a) ci-dessus ; ou (c) représentent des fragments, des variations alléliques ou autres d'une séquence d'ADN de la revendication 7, que ces variations résultent de changements dans la séquence peptidique ou non.
- 5 9. Séquence d'ADN suivant la revendication 8, qui est une séquence d'ADN génomique.
 - 10. Séquence d'ADN suivant la revendication 8, qui est une séquence d'ADNc.

35

40

45

50

55

11. Gène codant pour la BMP-4 humaine comprenant la séquence d'ADN suivante :

		p	turname compr	onant la ooquoi	IOC CAIDIT COIVE	1110 .		
10	30	20	70	40	50	. 60	70	
	CTCTAGAGGG	CAGAGGAGGA	OL DDADDDADDD	CAACCACAG	CCYCCOCCCC	cccesssects	GGTGAGTGTG	
15	80	90	100	110	120	<u>1</u> 30	140	
,,,	CCATOOTAGE	TGAGGGAGGC	CYCCLICYCY	CCCCTCCT	CCTCCCCCTG	AGTATCTAGC	TIGICICCC	
		160	170	180	190		210	
20	GATGGGATTC	COCIOCANCC	TATCTOGAGO	CIECHECECC	ACAGTCCCCG	CCCCTCCCCC	AGGITCACIG	
		230	240	250	260	270		
	CAACCITCA	executates	GGAGCTGCTG	CICCOCYCCC	CCCTACTGCA	GGGACCIALG	الماء الماء الماء	
25								
	290	300	310	320	330 ASCCTTTCCA	340		
	المتحافدتها	المماضينا	GLACIGUIGU	Mediature	ASCULLICCA	ــــــــــــــــــــــــــــــــــــــ		
	e er sammen		era e i e e e e e e e e e e e e e e e e e	Santan waka				. •
30	360 CIGICAAGAA	TCATCATCT	380 מתביית ביית	390	TGTCAAGACA	CC ATG ATT	∞	
		,				MET Ile	Pro	

5	GGT A GGT A	17 AC Sn	yrd CCY	ATG MET	CIG Leu	ATG MET	432 GIC Val	GIT Val	TTA Leu	TEA Leu	TGC CYS	447 CAA Gln	GIC Val	CIG Leu	CTA Leu	GGA GIY	462 GGC Gly	GCG Ala
	AGC C Ser H		CCT															
10	522 GGC C Gly H	AC is	GCG Ala	GGA GLY	CJA CCY	537 CCC Arg	OGC Arg	TCA Ser	GGG Gly	Gln Gln	552 AGC Ser	CAT His	GAG Glu	CIC Leu	CTG Leu	567 CGG Arg	GAC Asp	TTC Phe
15	GAG G	& :																
20	AGT G																	
	GAG G Glu G																	
25	AGC C Ser A	nd Re	GCC Ala	747 AAC Asn	ACC Thr	GTG Val	AGG Arg	AGC Ser	762 TTC Phe	CAC	CAC His	GAA Glu	GAA Glu	777 CAT His	CTG -Leu	GAG -Glu	AAC Asn	ATC
30	792 CCA G Pro G	:GG	AŒ	AGT	GAA	807 AAC	TCI	GCT	TTT	CI	822 TTC	crc	TTT	AAC	crc	837 AGC	AGC	ATC
35	oer o	AG ilu	852 AAC Asn	GJ <i>n</i> GAG	GIG Val	ATC Ile	TCC Ser	867 TCT Ser	GCA Ala	GAG Glu	CIT Leu	CGG Arg	882 CTC Leu	TTC Phe	CGG Arg	GAG Glu	CAG Gln	897 GTG Val
	GAC C	AG iln	G GGC	CCT Pro	GAT GAT ASp	TXP TXP	GAA Glu	AGG Arg	GGC Gly	927 TTC Phe	CAC	OGI Arg	ATA Ile	AAC Asn	942 ATT Ile	TAT Tyr	GAG Glu	GIT Val
	ATG A MET I	157 IAG Iys	∞ P±⊃	CCA Pro	GCA Ala	€£A €£AA	972 GIG Val	GIG Val	CCT Pro	GGG Gly	CAC His	987 CTC Leu	ATC	AĊA Thr	CGA Arg	CEA	1002 CTG Leu	yzb GyC
45	ACG A Thr A	rià cy	CIG	1017 GIC Val	CAC His	C-C Fis	AAT A <u>ŞD</u>	GIG	1032 ACA Thr	CGG Arg	TITE TIGG	GAA Glu	ACT	1047 TTT Phe	CAT ÇÇ	GTG Val	AGC Ser	CCI .Pro
50	1062 GCG G Ale V	TC	CTT	œc	TCC	1077 A∝	œc	ಚಾ	عدد	_C+G	1092 CCA	AAC	TAT	GGG	CTA	1107 G∝	. ATT	GAG .
55	GIG A Val I	CT	122 CAC His	CIC Leu	CAT His	eju ದೀಡಿ	ACT	1137 CGG Arg	ACC Thr	CAC His	CAG Gln	GGC	ll52 CAG Gln	CAT His	GTC Val	AGG Arg	ATT	1167 AGC Ser

		1182	1197	1	212
5	CCA TCG TTA Arg Ser Leu	CCT CAA GGG AGT	GGG AAT TGG	GCC CAG CTC CCG Ala Gln Ieu Arg	coc cre cre cre
	1227 ACC TIT GGC Thr Phe Gly	1242 CAT GAT GGC CGG His Asp Gly Arg	GCC CAT GCC	1257 TTG ACC CGA CGC Leu Thr Arg Arg	1272 CCG AGG GCC AAG Arg Arg Ala Lys
10	OGT AGC CCT	287 AAG CAT CAC TCA Lys His His Ser	1302 CAG CGG GCC Gln Arg Ala	1317 AGG AAG AAG AAT Arg Lys Lys Asn	AAG AAC TGC CCG Lys Asn Cys Arg
15	1332 CGC CAC TCI Arg His Ser	1347 CTC TAT GTG GAC Leu Tyr Val Asp	TTC AGC GAT	362 GIG GGC TGG AAT Val Gly Trp Asn	1377 GAC TGG ATT GTG Asp Trp Ile Val
20	1392 GCC CCA CCA Ala Pro Pro	GGC TAC CAG GGC	1407 TTC TAC TGC Phe TVr CVs	1422 CAT GGG GAC TGC His Glv Aso Cvs	1437 CCC TIT CCA CIG Pro Phe Pro Leu
				ATT GTG CAG AC	1482 CTG GTC AAT TCT Leu Val Asn Ser
25			A GOO TOT TO		1542 A CTG AGT GCC ATC 1 Leu Ser Ala Ile
30					7 A AAT TAT CAG GAG 5 ASN TYT Gln Glu
35	1602 AIG GIA GII MET Val Val	1617 A GAG GGA TGT GO 1 Glu Gly Cys Gl	(408) SG TGC CGC TG? .y Cys Arg	1636 10 AGATCAGG CAGTCCT	546 1656 TGA GGATAGACAG
	1666 ATATACACAC		1686 CACATAC ACCACA	1696 1706 ACACA CACGITCCCA	1716 1726 TOCACTCACC CACACCTAC
40	1736 ACAGACTGCT		1756 ACFITTA TTTAAJ		1786 1796 AATGGAAAAA ATCCCTAAAC
45		1816	1826 OTTIACS TGCAN	1836 1846	1856 1866 TCATCATATA TETTCACAAA
50		1886 AACEACSTAT TAA			1926 1936 TTAAAAAAA AAAAAACT
	1946 CIAGAGIOGA		·		
EE					

- 12. Gène codant pour la BMP-4 humaine comportant la séquence d'acides aminés donnée à la revendication 11.
- 13. Gène codant pour une protéine montrant des propriétés de la BMP-4 et comprenant une séquence d'ADN :

- (a) qui diffère d'une séquence d'ADN de la revendication 11 dans la séquence des codons du fait de la dégénérescence du code génétique ;
- (b) qui s'hybride avec une séquence d'ADN de la revendication 11 ou du paragraphe (a) ci-dessus ; ou
- (c) représente un fragment, une variation allélique ou autre d'une séquence d'ADN de la revendication 11, que cette variation résulte de changements dans la séquence peptidique ou non.
- 14. Séquence d'ADN suivant la revendication 13, qui est une séquence d'ADN génomique.
- 15. Séquence d'ADN suivant la revendication 13, qui est une séquence d'ADNc.
- 16. Vecteur contenant le gène ou la séquence d'ADN suivant l'une quelconque des revendications 1 à 15, en association active avec une séquence de contrôle d'expression.
- 17. Cellule transformée avec un vecteur de la revendication 16.

5

10

15

20

30

40

45

- 18. Cellule suivant la revendication 17, qui est une cellule mammifère, une cellule bactérienne, une cellule d'insecte ou une cellule de levure.
- 19. Cellule suivant la revendication 18, qui est une cellule CHO.
- 20. Protéine montrant des propriétés de la BMP-2, qui est codée par un gène ou une séquence d'ADN de l'une quelconque des revendications 1 à 10.
- 21. Protéine montrant des propriétés de la BMP-2, qui est obtenable par les étapes de culture dans un milieu de culture approprié d'une cellule transformée avec un vecteur d'expression comprenant un gène ou une séquence d'ADN de l'une quelconque des revendications 1 à 10, et de récupération de ladite protéine du milieu de culture précité.
 - 22. Protéine montrant des propriétés de la BMP-4, qui est codée par un gène ou une séquence d'ADN de l'une quelconque des revendications 11 à 15.
 - 23. Protéine montrant des propriétés de la BMP-4, qui est obtenable par les étapes de culture dans un milieu de culture approprié d'une cellule transformée avec un vecteur d'expression comprenant un gène ou une séquence d'ADN de l'une quelconque des revendications 11 à 15, et de récupération de ladite protéine du milieu de culture précité.
- 24. Procédé de production de la protéine suivant l'une ou l'autre des revendications 21 et 23, comprenant les étapes de culture dans un milieu de culture approprié de la cellule de la revendication 17 et d'isolement de ladite protéine du milieu de culture précité.
- 25. Composition pharmaceutique comprenant les protéines de l'une quelconque des revendications 20 à 23, individuellement ou en combinaison, et un véhicule pharmaceutiquement acceptable.
 - 26. Composition pharmaceutique suivant la revendication 25, comprenant de plus une matrice pouvant distribuer la composition au site de l'anomalie osseuse ou cartilagineuse et formant une structure pour induire une formation osseuse ou cartilagineuse.
 - 27. Composition pharmaceutique suivant la revendication 26, dans laquelle ladite matrice comprend de l'hydroxyapatite, du collagène, de l'acide polylactique ou du phosphate tricalcique.
- 28. Utilisation d'une protéine suivant l'une quelconque des revendications 20 à 23, individuellement ou en combinaison, pour la préparation d'une composition pharmaceutique pour induire une formation osseuse ou cartilagineuse.

Revendications pour l'Etat contractant suivant : AT

Procédé de préparation d'un gène codant pour la BMP-2 humaine comprenant la séquence d'ADN suivante :

	cica		10 TA (EAGT(3T C	AGCA(3C CITGO		GGG	.40 ACIT	CTT	SAACI	50 MG (CAGGO		00 A T.	ACTTO	70 CCCA
5	œ		80 TG (œ		90 X I	TIGO	100 XX		SAGO	110	TTO	_	120 TCT	00GA(30 ZA C		140 ETCC
10	ACTO		.50 :GC (CTTGC		50 AC A	.CTGA	170 34030		rtœ	180 2460	GTG		L90 SAG	AGAC.		00 33 G	cosso	210 ACCC
15	GGGA		20 GA (GAG(¥G ≯	እእእርኛ	240 =\ac		TEAT	250 36GT	CCT		260 CCA (GGIC		70 ' EA C	CACAC	280 TTTT
20	TOCA		9,0 GA (ŒCIY		00. 2A A	TGEAC	310 ŒIGI			320 31GC	TTĊ		330 406 (GACTO	34 30067		CCIA	:350 AGGT
	<u>αε</u> λα	(1) AT ME	ច្ច	ig gg	x cc la Gl	IG A	70 CC CC	:3 c) :c 10	n C	T C.	EA GO	35 X T la Le	er it	rg C	TT CO BU Pr	$\propto c$	in v	TC el	
25				415					430					445					
	CTC (Leu I	eu (GGC Gly	GC	GCG Ala	GCT Ala	GJA GGC	CTC Leu	GTT Val	CCG Pro	GAG Glu	CTG Leu	GGC Gly	œc	AGG	Tā2 Yyc	TTC Phe	SCC	
30	460 GCG (Ala /						∞												
			520			•		535					550					565	
35	Phe (
	ASG (خت 375	GCC Ala	GTG Val	580 GTG Val	α 220		TAC TYT	ATG MET	595 CTA Leu	у г ь С7С	CTG Leu	T-T Tyr	yrg ŒC	yrd ycc 910	CAC His	TCS Sex	GIY GIY	
40	,	625					640					655					670		
	CAG (∞	GGC Gly	TCA Ser	000 Pro	GCC Ala	CCA	S G G A C	CAC His	ಜ್ ಜ್	TTG Leu	GYC.	Arg Arg	GCA Ala	GCC Ala	AGC Ser	CCA.	GCC	
45	AAC A	ACT Ihr	GTG Val	685 CGC	AGC Ser	TIC	CAC	CAT	700 <u>G2A</u> Glu	€jπ € j ÿ	TCT Ser	TTG Leu	GAA Glu	715 GAA Glu	CIA	CCA Pro	Glu	, ACG I Thr	
				_															

	730)				745					760	l.				775		
	AGI	, ecc	AAA	ACA	λœ			عين .	كين	للبليلة			201	بليك	ATC	~~~	λŒ	GAG
	Ser	Gly	Lys	Tur	Th	Ard	Ara	Phe	Phe	Phe	Asn	Leu	Ser	Ser	Ile	Pro	Thr	Glu
5						_	,											
			790				•	805					820					835
	GAG	TII	' ATC	: ACC	TCA	GCA	GYC.	CIT	CAG	GIT	TIC	Œλ	CAA	CYC	ATG	CAA	CAT	GCT
	GIU	Phe	: He	Thr	Ser	Ala	Glu	Ieu	Gln	Val	Phe	Arg	Glu	Gln	MET	Gln	Asp	Ala
					250													
10	בידיני	CC3	ממ	יייבגי	850		mm-	~~		865					880			
	Leu	Glv	Asn	TAA : nzA :	Sor	Sor	Tic	71.	uac wie	7	ATT	AAT	ATT	TAT	GAA	AIC	ATA	AAA
						Jer	FIRE	nıs	RIS	Arg	Пe	ASI	тте	ıvr	Giu	TTG	тте	тХг
		895					910					925					940	
	∞ T	GCA	ACA	. 600	AAC	TOG		TTC	∞	GTG	ACC		CIT	TTG	GAC	ACC	AGG	TTG
15	Pro	Ala	Thr	Ala	Asn	Ser	Lys	Phe	Pro	Val	Thr	Ser	Leu	Leu	ASD	Thr	Aru	Ieu
							•										5	
				955					970					985				
	GIG	TAA	CAG	AAT	GCA	AGC	AGG	TGG	CYY	AGT	TTT	GAT	GTC	ACC	∞	GCT	GIG	ATG
20	٧٤١	AST	GID	Asn	Ala	Ser	Arg	متت	GJn	Ser	Phe	Asp	Val	Thr	Pro	Ala	Val	MET
	100	^																
			a cm	-		1015	~				1030					1045		
	Are	ענען.	سدائيل برمود	GCA Ala	Cla	Clir	271.0	315	AAC	CAT	GGA	TIC	GIG	GIG	CAA	GIG	GCC	CAC
				. Ma	911	GLY.	LTZ	WIG	WZI	ETZ	GTĀ	Prie	٧ڪ	var	GIU	val	ALZ	-115
25			1060				-	1075					1090					1105
	TTG	C+C	GAG	AAA	CAA	.GGT			AAG	AGA	CAT	GIT	AGG	ATA	AGC	AGG	TCT	TIG
	' Leu	Glu	Glu	Lys	Gln	Gly	Val	Ser	Lys	Arg	His	Val.	'Airg	Ile	ser	Arg	ser	<u>.eu</u>
													_			_	-	
	CSC	~:	C) M		1750				:	1135					1150			
30	#ic	Cln	gw _T	GAA	CAC	AGC	TGG	TCA	CAG.	ATA	AGG	\overline{x}	TIG	CIA	GIA	ACT	TIT	GGC
	نسنه ه	GIII	ນວັກ	Glu	RIS	ser	JID	Ser	Gin	тте	Arg	pro	Leu	Teu	Val	Tr	me	GIY
	1	1165				1	180				,	1195					210	
	CAT	GAT	GGA	AAA	GGG			crc	CAC	AÁA	אכז	727	222	ىلىن	C22	ഹ്	222	Cac
35	عنة	λsp	Gly	Lys	Gly	His	Pro	Ieu	His	Lvs	Ara	Glu	INS	Aru	Gln	Ala	Ivs	His
				_	•						,			5				
				1225					240				1	.255				
	AAA T	CAG	æ	AAA	α C	CTT	AAG	∞	AGC	TGT	24K	AGA	CAC	α	TTG	TAC	GTG	GYC
	TÀR	GTU	Arg	Γλ2	Arg	Leu	Lys	Ser	Ser	Cys	Lys	Arg	His	ಮಾ	Leu	Tyr	Val	ಭಿತ್ರಾ
40	1270	3			,	205												
			G2C	GIG		285	2200	C) C	<u>ب</u>	7	300		~~	~~		315	~~	~~
	Phe	Ser	Asp	Val	Glv	لمحر <u>ل</u> 1997	7 <i>2</i> 2	700	سمن ميتر _آ	710	7/2)	332	2~~	D		17-47.	Wie -	212
		_	. 25	•	GIŢ	لوخد	P-311	بتت	ليند	TIE	٠	Mic	PiJ	210	GTĀ	IĀŢ	حدہ	7 16
		2	330				1	345				٦	350				٦	375
45	$\overline{\Lambda}\overline{\Lambda}\overline{\Lambda}$	TAC	TGC	CAC	GGA	CAA			TIT	α	CIG			CAT	CIG	AAC		
	Phe	Ty	Cys	His	Gly	Glu	Cys	Pro	Phe	Pro	Leu	Ala	حجد	His	Leu	Asn	Ser	<u> </u>
					-		-						•					•
		— –			.390					405				1	420			
50	AAT	CYI.	∞	ATT	CIT	CAG .	ACG '	TIG	CIC	AAC	TCT	GIT .	AAC	TCT	AAS.	TIA	α	عدد
50 .	ຂຸຂກ	عدۃ	Ala	TIP	Val	Cln '	Thr -	Tou	(c \	300	Sar	1/a 1	752	C2-	Tue	TIO	لتحظ	Tyc

1435 1450 1465 1480 GCA TGC TGT GTC COG ACA GAA CTC AGT GCT ATC TCG ATG CTG TAC CTT GAC GAG Ala Cys Cys Val Pro Thr Glu Leu Ser Ala Ile Ser MET Leu Tyr Leu Asp Glu 5 1495 1510 1525 AAT GAA AAG GIT GIA TIA AAG AAC TAT CAG GAC AIG GIT GIG GAG GGI TGI GGG Asn Glu Lys Val Val Leu Lys Asn Tyr Gln Asp MET Val Val Glu Gly Cys Gly 10 1540(396) 1553 1563 1573 1583 1603 TGT CSC TAGTACAGCA AAATTAAATA CATAAATATA TATATATATA TATATITTIAG AAAAAAGAAA Cys Arg

15 AAAA,

20

30

35

45

50

dans lequel ledit procédé comprend les étapes suivantes

- a) la sélection d'une bibliothèque de gènes construite à partir d'ADN ou d'ADNc provenant de U-2 OS avec un fragment de bBMP-2 marqué par hybridation,
- b) l'isolement des clones positifs, et
- c) l'isolement des inserts d'ADN de ces clones.
- Procédé suivant la revendication 1, dans lequel le gène code pour la BMP-2 humaine ayant la séquence d'acides aminés donnée à la revendication 1.
 - 3. Procédé de préparation d'un gène codant pour une protéine montrant des propriétés de la BMP-2 humaine et comprenant une séquence d'ADN:
 - a) qui diffère d'une séquence d'ADN de la revendication 1 dans la séquence des codons du fait de la dégénérescence du code génétique;
 - b) qui s'hybride avec une séquence d'ADN de la revendication 1 ou du paragraphe (a) ci-dessus ; ou
 - c) représente un fragment, une variation allélique ou autre d'une séquence d'ADN de la revendication 1, que cette variation résulte de changements dans la séquence peptidique ou non,

dans lequel le procédé susdit comprend des techniques standards de biologie moléculaire.

- 4. Procédé suivant la revendication 3, dans lequel la séquence d'ADN est une séquence d'ADN génomique.
- Procédé suivant la revendication 3, dans lequel la séquence d'ADN est une séquence d'ADNc.
 - 6. Procédé de préparation d'un gène codant pour la BMP-2 bovine comprenant la séquence d'ADN suivante :

```
45
                                          30
                       15
      GGC CAC GAT GGG AAA GGA CAC CCT CTC CAC AGA AGA GAA AAG CGG
                                                         X
                                                             R
                                             3
                                                 R
                                                     Ξ
                                 P .- L
                                         H
                             H
                          Ğ
5
                                                              90
                                          75
      CAA GCA AAA CAC AAA CAG CGG AAA CGC CTC AAG TCC AGC TGT AAG
                                                             X
                                                     S
                                             K
                                                 S
                 H
                      X
                          Q
                              R
                                 K
                                     R
10
                                          120
                      105
      AGA CAC CCT TTA TAT GTG GAC TTC AGT GAT GTG GGG TGG AAT GAC
                                 F
                                         ם
                                             v
                                                 G
                                                     W
                          v
                              D
                                     S
                                                             180
                                          165
                      150
      TGG ATC GTT GCA CCG CCG GGG TAT CAT GCC TTT TAC TGC CAT
                                                             GGG
15
                                              F . Y
                                     H
                          P
                              G
                                 Y
                                          A
                                                              225
                                         .210
                      195
      GAG TGC CCT TTT CCC CTG GCC GAT CAC CTT AAC TCC ACG AAT CAT
                                                             H
                              A D
                                             N
                                                 S
                                                     T N
                                     H L
                          L
20
                      Ъ
                                                              270
                                          255
                      240
      GCC ATT CTC CAA ACT CTG GTC AAC TCA GTT AAC TCT AAG ATT CCC
                                                              5
                                                      K
                                                         I
                              V
                                  N
                                      S
                                              N
                                                  S
              v
                          L
                  Q
                      T
25
                                                              315
                                          300
                      385
      AAG GCA TGC TGT GTC CCA ACA GAG CTC AGC GCC ATC TCC ATG CTG
                                                  I
                                                      S
                                                         М
                              T
                                  Ξ
                                          S
                                              A
                                                              360
                                          345
                      330
30
      TAC CTT GAT GAG AAT GAG AAG GTG GTA TTA AAG AAC TAT CAG GAC
                                                              ס
                                              К
                                                  И
                              К
                                          L
                      N
                          Ε
                                  V
                                                            407
                                                 397
                      375 .
                                      (129)
      35
                                      R
                  Ξ__G
                          C \cdot G
                                  С
                                                              457
                                                  447
                417
                            427
                                       437
40
         TAAATATATA TATATATA TTAGAAAAAC AGCAAAAAAA TCAAGTTGAC
                                                              507
                                       487
                                                  497
                467
                            477
         ACTITAATAT TTCCCAATGA AGACTITATT TATGGAATGG AATGGAGAAA
45
                            527
                                                              557
                                       537
                                                  547
         AAGAAAACA CAGCTATTTT GAAAACTATA TTTATATCTA CCGAAAAGAA
                567
                            577
                                       587
         GTTGGGAAAA CAAATATTTT AATCAGAGAA TTATT,
50
```

dans lequel ledit procédé comprend les étapes suivantes :

a) la sélection d'une bibliothèque de gènes construite à partir d'ADN ou d'ADNc provenant de foie bovin avec une sonde marquée conçue sur la base de la séquence d'acides aminés d'un fragment de bBMP-2,

b) l'isolement des clones positifs, et

c) l'isolement des inserts d'ADN de ces clones.

^{7.} Procédé suivant la revendication 6, dans lequel le gène code pour de la BMP-2 bovine ayant la séquence d'acides

aminés de la revendication 6.

8. Procédé de préparation d'un gène codant pour une protéine montrant des propriétés de la BMP-2 bovine et comprenant des séquences d'ADN: 5 a) qui diffèrent d'une séquence d'ADN de la revendication 7 dans la séquence des codons du fait de la dégénérescence du code génétique; b) qui s'hybrident avec une séquence d'ADN de la revendication 7 ou du paragraphe a) ci-dessus ; ou c) représentent des fragments, des variations alléliques ou autres d'une séquence d'ADN de la revendication 10 7, que ces variations résultent de changements dans la séquence peptidique ou non, dans lequel le procédé précité comprend des techniques standards de biologie moléculaire. 9. Procédé suivant la revendication 8, dans lequel la séquence d'ADN est une séquence d'ADN génomique. 15 10. Procédé suivant la revendication 8, dans lequel la séquence d'ADN est une séquence d'ADNc. 11. Procédé de préparation d'un gène codant pour la BMP-4 humaine comprenant la séquence d'ADN suivante : 20 25

35

40

45

50

	10 CTCTAGAGGG	CACACCACCA	GGGAGGGAGG		50 GCAGCCCGC	œê₹ÆŒ¥ €0	70 GGTGAGTGTG
5	08 CONTOCOLACO	90 TGAGGGAGGC	100 GAGOCTGAGA				140 TIGICICCC
10	150 GATGGGATIC	160 CCTCCAACC	170 TATCTOGAGO	180 CICCAEGGC	190 ACAGTCCCCG	200 CCCTCCCCC	210 AGGITCACTG
15	220 CAACCGTTCA	230 EAGTOOGA	240 GGAGCTGCTG	250 CTGGCCAGCC	260 CCTACTGCA	270 GGGACCTATG	280 GAGCCATICC
	290 GIAGIGCCAT	300 300	310 GCACTGCTGC	320 ACCITOCCIG	330 AGCCTTTCCA	340 ÇCAAGITIGI	350 TCPACATTGG
20	360 CTGTCAAGAA	370 TCATCCACIG		390 CITGITITC	400 TGTCAAGACA	\-/	
25							
30	417 GGT AAC C Gly Asn A	rd hei Ten ry Yie Cie	432 ATG GTC GTT MET Val Val	TIA TIA TO Leu Leu Cy	447 SC CAA GTC C TS Glm Val I	TG CTA GGA Leu Leu Gly	462 GGC GCG Gly Ala
	AGC CAT (Ser His A	477 SCT AGT TIG Lla Ser Leu	ATA CCT GAO Ile Pro Glu	492 ACG GGG AA Thr Gly Ly	C AAA AAA Q	507 STC GCC GAG /al Ala Glu	ATT CAG Ile Gln
35	522 GGC CAC O Gly His A	rja ejh ejh ece ech ech	537 CCC CCC TCP Arg Arg Ser	55 A GGG CAG AC Gly Gln Se	C CAT CAG (557 CTC CTG CCG Leu Leu Arg	GAC TTC Asp Phe
40	ಆಎ ಯ ಸ	582 ACA CIT CIG Bur Leu Leu	597 CAS AIG TIT Gln MIT Phe	ന്ദ്രേയ വാര്യ	g Arg Arg 1 612	cae cae cae	627 AGC AAG Ser Lys
45	AGT GCC (Ser Ala V	642 FIC ATT COS /al Ile Pro	GAC TAC ATO Asp Tyt MET	657 CGG GAT CI Arg Asp Le	T TAC CGG (au Tyr Arg I	672 TIT CAG TCT Leu Gln Ser	ece ere
	687 GAG GAG G Glu Glu G	EAA GAG CAG Slu Glu Gln	702 ATC CAC AG Ile His Ser	ACT GGT CI	717 TT GAG TAT (eu Glu Tyr I	CT GAG CGC Pro Glu Arg	732 CCG GCC Pro Ala
50							·

				747					762					777				
												€ 7 7						
	Ser	Arg	Ala	Asn	The	Val	yrg	Ser	Phe	عنة	His	Glu	Glu	His	-Leu	-Glu	Asn	Ile
5														•				
	792	~~	3~			807					822					837		
												CIC						
	0	GTJ	1111	SET	GIU	MS()	Ser	WIG	File	wig	File	Leu	P.IE	W-Zr I	LEU	عحت	عجد	776
10			852					867	•				882					897
10	α T	CYC		GAG	GTG	ATC	TO		GCA	GAG	CIT	œ		TTC	œ	GAG	CAG	
												Arg						
												_				•		
	~ ~				912					927					942			
15												ŒĨ						
	wzņ	GIN	CTA	PIO	vzĎ	лъ	GIU	Arg	GTĀ	me	HIS	yrg	щe	ASTI	TTE	JÄL	GIU	AST
		957					972					987				,	1002	
	ATG		∞	CCA.	GCA	GAA		GTG	CIT	GGG	CAC	crc	ATC	ACA	ŒĄ			CAC
	MET	Lys	Pro	Pro	Ala	Glu	Val	Val	Pro	Gly	Fis	Leu	Ile	Thr	Arg	Leu	Leu	ASO
20		-								-								-
			-	.017					1032					L047				
												G⊅Y.						
	THE	vid.	Leu	Val	riis .	His	YZU	.Val	Thr.	ynd'	متن	Glu	Inr	Phe	yżb	.Val	Ser	Pro
25	1052	,			1	.077				,	.092				,	1107		
			CIT	œc			œ	GAG	75G			220	ידביד	ccc			ידינים.	GAG
	Ala	vai	Leu	Aru	Tim.		Ar u	Glu	LVS	Gln	<u> </u>	Asn	īvr	Glv.	Leu	Ala	Ile	Glu
				•	•		•		•					-				
		_	122					.137					.152					167
30	GIG	ACT	CAC	CTC	CAT	ಧ್ಯ	ACT	œ	∞	CYC	CYC	GGC	CYC	CAT	GIC	AGG	ATT.	AGC
	Val	mr	HIS	Ten	215	GIN	Thr	yrg	Inr	715	Gln	Gly	Gln	His	Val	Arg	Ile	Ser
				3	100				,	107				-				
35	œa.	TOT	בחה		182		الت) د	ccc		197	~~	CAG	حلاء		1212	\sim	CTG	9
	yrg																	
	-					1		1										
	•	227					242					257				-	272	
												∞						
40	Thr	Phe	Gly	His	Αsp	Gly	Arg	Gly	His	Ala	Leu	Thr	Arg	Arg	Arg	Arg	Ala	Lys
			٦	287				,	.302				-	L317				
	ŒT	AGC			ראַיי	CAC	ערש			GCC	agg	AAG			226	AAC	TGC	CGG
												Lys						
45				-					•		-	-	•		•		-	_
40	1332					347					L362					1377		
												GGC						
	Arg	HIS	Ser	Leu	TYT	Val	تحز	Phe	Ser	تجر	Val	Gly	dxL	Asn	٧ڪ	J.J.	TTS	VZI
		. 1	L392				٠,	407					1422					1437
50	GCC			GGC	T2.	CG G			חבר	TCC	CP.T	GGG		TCC	ccc.	TIT		
												Glv						
	~~	m ~:	<i>-</i>		145 ~ ` ` `	2 ~~		~ · ·	c	146		بہ س	Y (1)	C 30	148 ص ح		γ 22	ىلىكىت بات
	GC 3.1	مد د حصالت	יים ביי אם ביי	CI	د د د.	C IC	A AC	بر المراجد - مراجد	n Hi	ىد ك د 14 د	~ Al 71	⊥ U.⊐	3 C1	ከ ሞ		nı Va	1 As	T TCT n Ser
55	ئىس.	a AS	ה עו	عدد	urs	" >S	لندين	~ ~~	a i Fil	ــ مــ	طند خه.	_	- 61					

GTC AAT TOO AGT ATC COO AAA GOO TGT TGT GTG COO ACT GAA CTG AGT GOO ATC Val Asn Ser Ser Ile Pro Lys Ala Cys Cys Val Pro Thr Glu Leu Ser Ala Ile TOO ATG CTG TAC CTG GAT GAG TAT GAT AAG GTG GTA CTG AAA AAT TAT CAG GAG Ser MET Leu Tyr Leu Asp Glu Tyr Asp Lys Val Val Leu Lys Asn Tyr Gln Glu (408)ATE GTA GTA GAG GGA TGT GGG TGC CGC TGAGATCAGG CAGTCCTTGA GGATAGACAG MET Val Val Glu Gly Cys Gly Cys Arg ATATACACAC CACACACACA CACCACATAC ACCACACA CACCTTOCCA TOCACTCACC CACACACTAC ATTCACCITE ACCITATE TCACITIACO TECANATETT TICACCATAT TCATCATATA TTTTCACANA: ATATATTTAT AACTAGGTAT TAAAAGAAAA AAATAAAATG AGTCATTATT TTAAAAAAAA AAAAAAAACT CTAGAGTOGA CGGAATTC, dans lequel le procédé précité comprend les étapes suivantes : a) la sélection d'une bibliothèque de gènes construite à partir d'ADN ou d'ADNc provenant d'U-2 OS avec un fragment bBMP-2 marqué par hybridation, b) l'isolement des clones positifs, et c) l'isolement des inserts d'ADN de ces clones. 12. Procédé suivant la revendication 11, dans lequel le gène code pour la BMP-4 humaine ayant la séquence d'acides aminés donnée à la revendication 11. 13. Procédé de préparation d'un gène codant pour une protéine montrant des propriétés de la BMP-4 et comprenant une séquence d'ADN: a) qui diffère d'une séquence d'ADN de la revendication 11 dans la séquence des codons du fait de la dégénérescence du code génétique : b) qui s'hybride avec une séquence d'ADN de la revendication 11 ou du paragraphe a) ci-dessus ; ou c) représente un fragment, une variation allélique ou autre d'une séquence d'ADN de la revendication 11, que cette variation résulte de changements dans la séquence peptidique ou non,

dans lequel le procédé précité comprend des techniques standards de biologie moléculaire.

- 14. Procédé suivant la revendication 13, dans lequel la séquence d'ADN est une séquence d'ADN génomique.
- 15. Procédé suivant la revendication 13, dans lequel la séquence d'ADN est une séquence d'ADNc.
- 16. Vecteur contenant le gène ou la séquence d'ADN préparé suivant l'une quelconque des revendications 1 à 15, en

association active avec une séquence de contrôle d'expression.

- 17. Cellule transformée avec un vecteur de la revendication 16.
- 5 18. Cellule suivant la revendication 17, qui est une cellule mammifère, une cellule bactérienne, une cellule d'insecte ou une cellule de levure.
 - 19. Cellule suivant la revendication 18, qui est une cellule CHO.
- 20. Procédé de préparation d'une protéine montrant des propriétés de la BMP-2, dans lequel ledit procédé comprend les étapes de culture dans un milieu de culture approprié d'une cellule transformée avec un vecteur d'expression comprenant un gène ou une séquence d'ADN préparé suivant l'une quelconque des revendications 1 à 10 et de récupération de ladite protéine du milieu de culture précité.
- 21. Procédé de préparation d'une protéine montrant des propriétés de la BMP-4, dans lequel ledit procédé comprend les étapes de culture dans un milieu de culture approprié d'une cellule transformée avec un vecteur d'expression comprenant un gène ou une séquence d'ADN préparé suivant l'une quelconque des revendications 11 à 15 et de récupération de ladite protéine du milieu de culture précité.
- 20. Procédé de production d'une protéine montrant des propriétés de la BMP-2 ou BMP-4, comprenant les étapes de culture dans un milieu de culture approprié de la cellule de la revendication 17 et d'isolement de ladite protéine du milieu de culture précité.
- 23. Procédé de préparation d'une composition pharmaceutique comprenant la combinaison des protéines préparées suivant l'une quelconque des revendications 20 à 22, individuellement ou en combinaison avec un véhicule pharmaceutiquement acceptable.
 - 24. Procédé suivant la revendication 23, dans lequel la composition pharmaceutique susdite comprend de plus une matrice pouvant distribuer la composition au site de l'anomalie osseuse ou cartilagineuse et constituer une structure pour induire une formation osseuse ou cartilagineuse.
 - 25. Procédé suivant la revendication 24, dans lequel la matrice comprend de l'hydroxyapatite, du collagène, de l'acide polylactique ou du phosphate tricalcique.
- 26. Utilisation d'une protéine préparée suivant l'une quelconque des revendications 20 à 22, individuellement ou en combinaison, pour la préparation d'une composition pharmaceutique pour induire une formation osseuse ou cartilagineuse.

55

30

40

45